

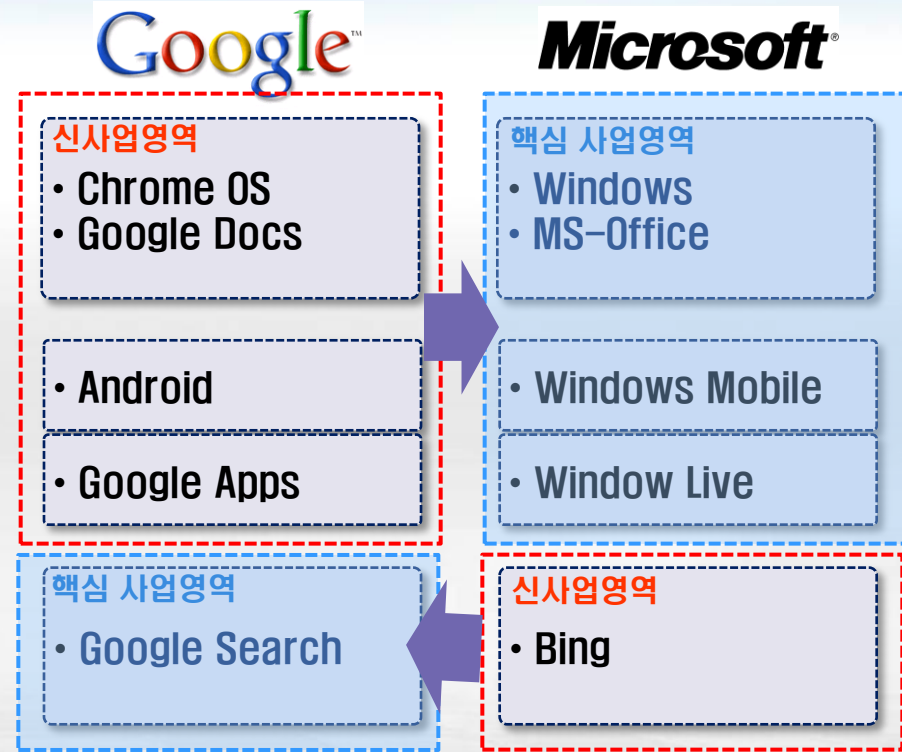
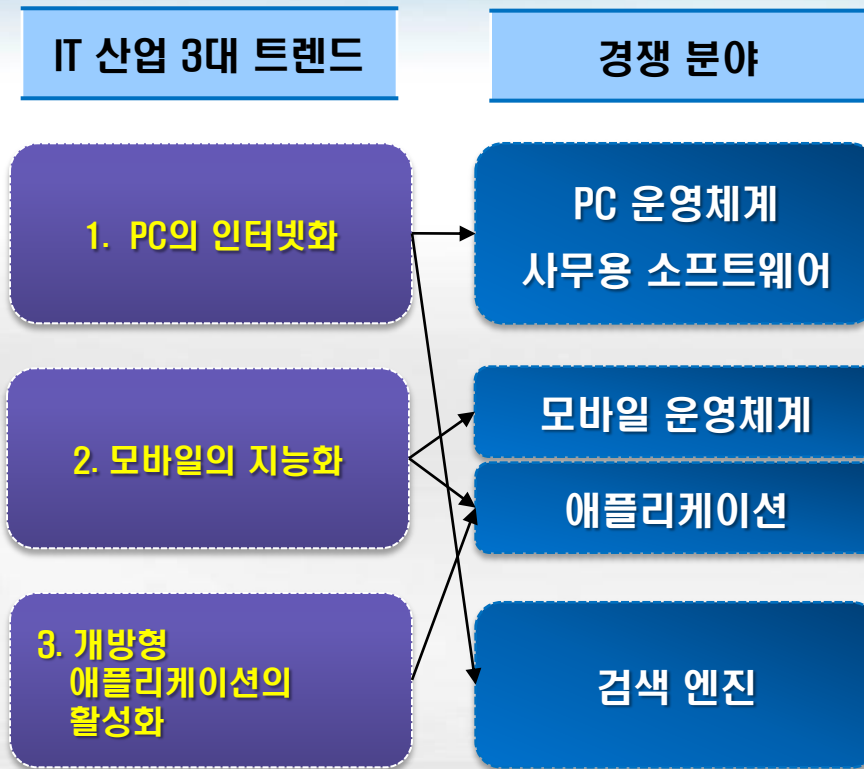
클라우드 컴퓨팅 기반 스마트그리드 데이터 분석 플랫폼 개발

2010. 7. 21

SK C&C 클라우드컴퓨팅사업본부
이 덕 재 부장

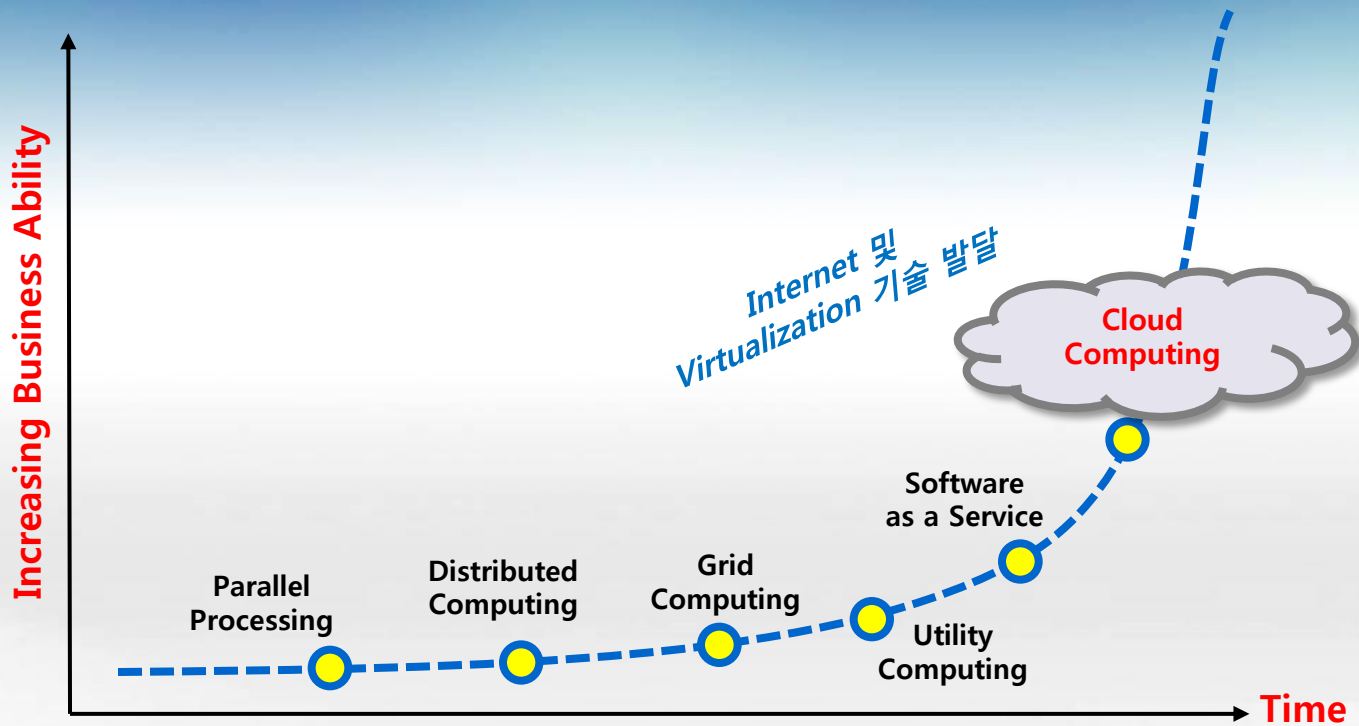
클라우드 컴퓨팅의 이해

IT 산업의 3대 트렌드



(Source : SERI 경영노트 20호, 2009.8.20)

Computing 관련 기술의 진화

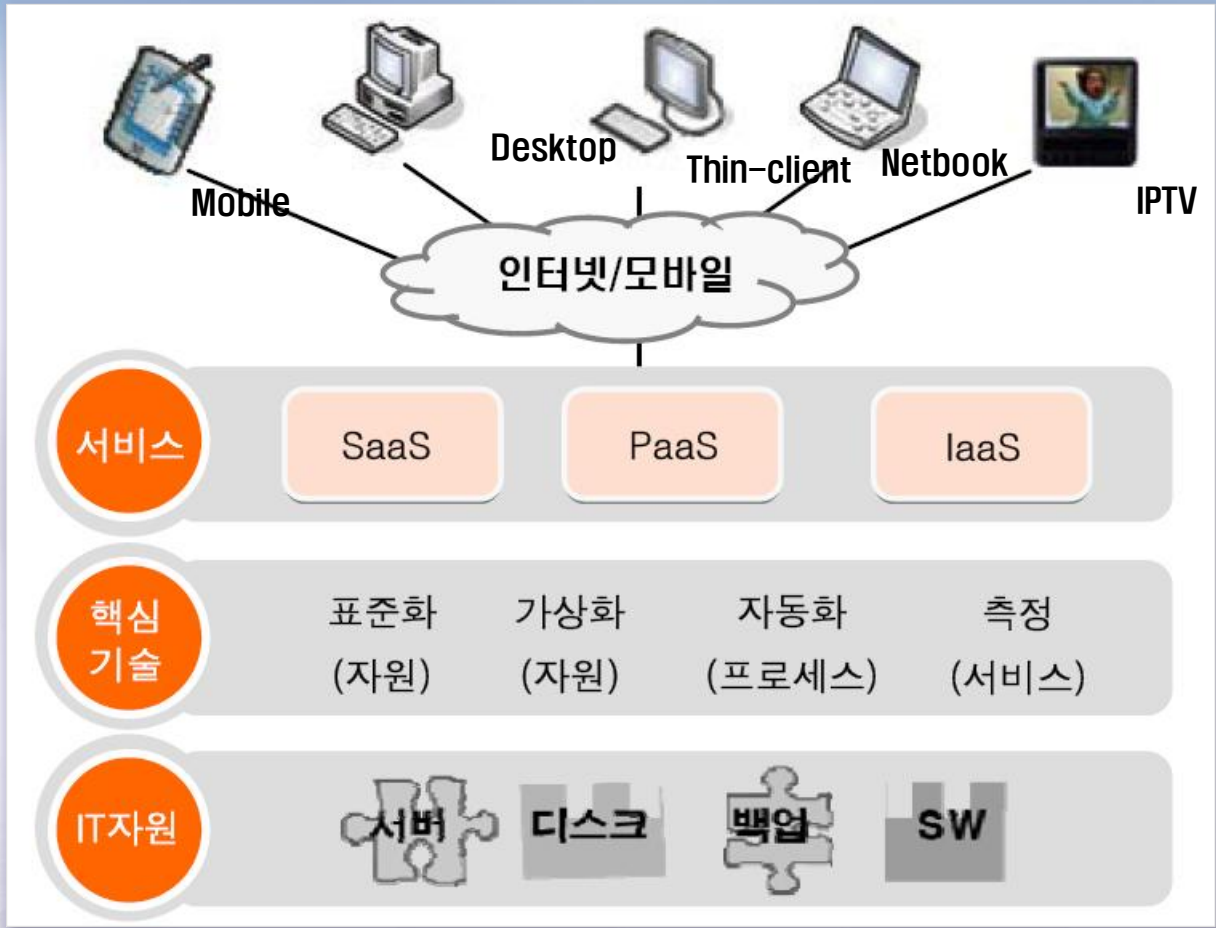


과거 관련 Computing 기술의 Mainstream 형성 실패

- 표준 플랫폼화 실패 (범용성 부족)
- 복잡하고 이질적인 컴퓨팅 환경 수용 시도 (특정 벤더 종속 등)

- 인터넷 접속 환경
- 가상화 기술 발전
- x86 서버 성능 향상

Cloud Computing 개념도



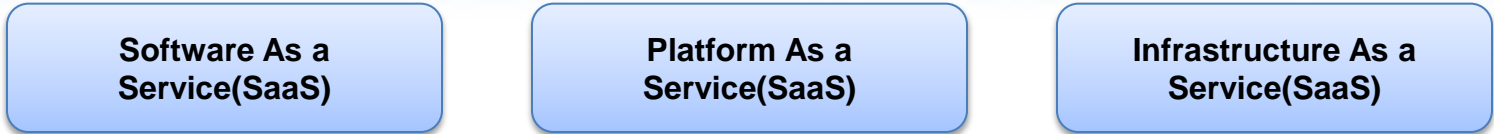
- SaaS : Software as a Service
- PaaS : Platform as a Service
- IaaS : Infrastructure as a Service

Cloud Computing Stack

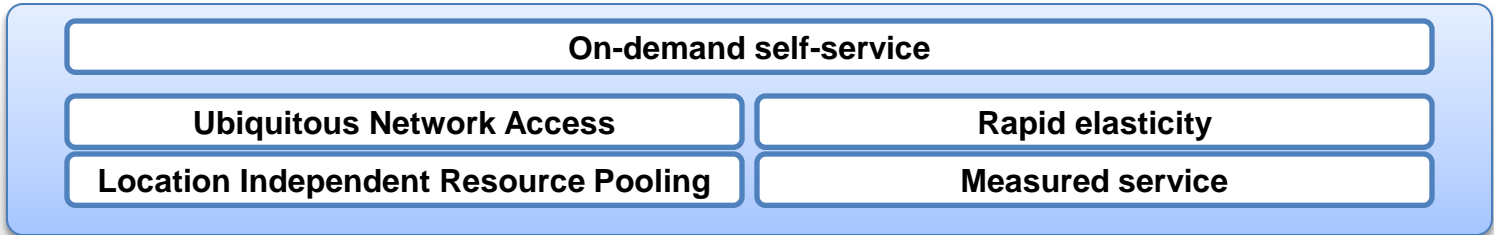
Deployment Models



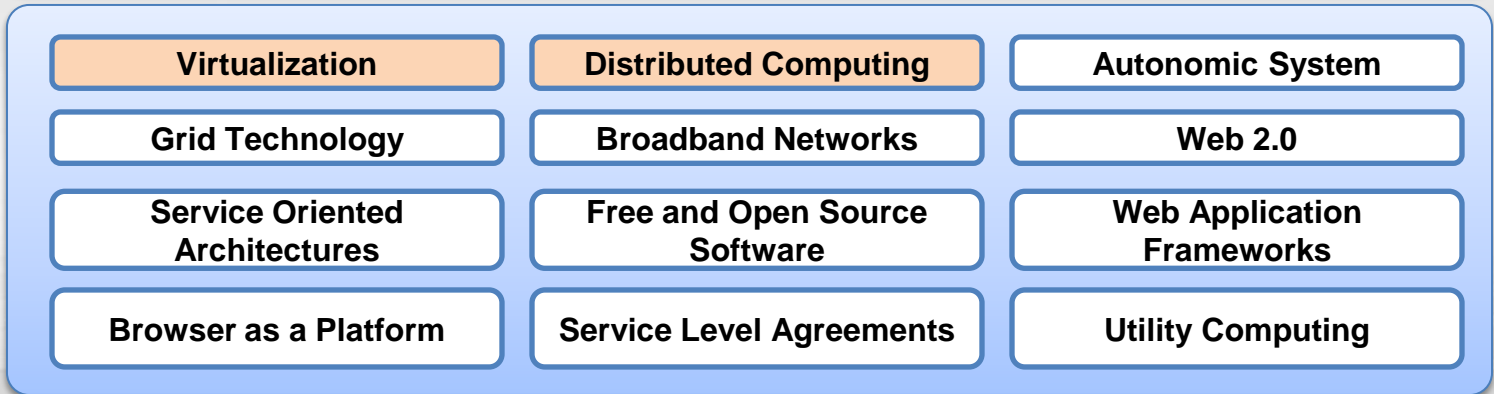
Delivery Models



Essential Characteristics



Foundational Elements/Enablers

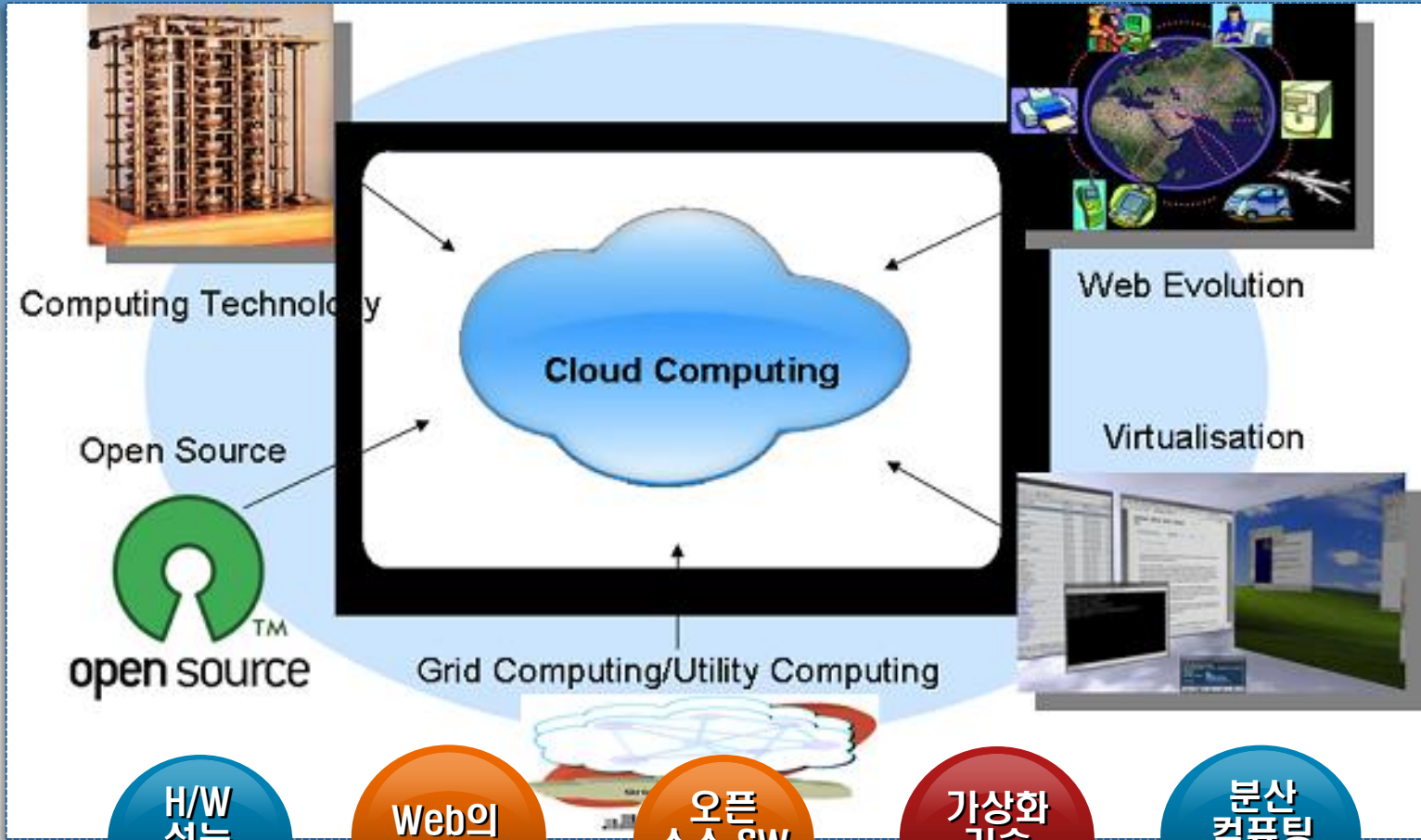


(Source: Working Definition of Cloud Computing from National Institute of Technology)

Cloud Computing 주요 기술 (공개SW)

구분	기술 분류	공개 소프트웨어
가상화	서버 가상화	• KVM, Xen, VirtualBox, Libvirt
	J2EE 3tier	• Apache Web, Tomcat, JBoss, MySQL, PostgreSQL
	개발 Framework	• Spring, Struts, IBatis, Hibernate 등
	Provisioning	• OpenNebula, RHEV, OVirt
분산 컴퓨팅	클라우드 스토리지	• Hadoop DFS, MogileFS, GridFS, GlusterFS
	분산 병렬 처리	• Hadoop MapReduce, openMPI, Hive, ZooKeeper
	분산 DB	• Neptune, MongoDB, Cassandra, GreenPlum, Memcached
	데이터 마이닝 프레임워크	• Wecka, RapidMiner, Mahout
클러스터 관리	모니터링	• Ganglia, Zenos, Chukwa, Hyperic
	리소스 스케줄러	• Spring Batch, ZooKeeper, Condor
	메시지 큐	• Qpid, ActiveMQ

Cloud Computing Drivers



H/W
성능
향상

Web의
진화

오픈
소스 SW
확산

가상화
기술
발전

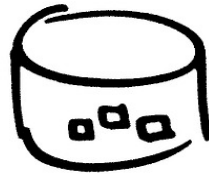
분산
컴퓨팅
기술

Cloud Computing 핵심 기술 – Hadoop (1)

Hadoop – 대용량 분산 데이터 저장 및 처리 시스템 (Big Data + Analytics)



Commodity Hardware



More Scalable than a Database / RDBMS



Methods Developed by Google



Apache

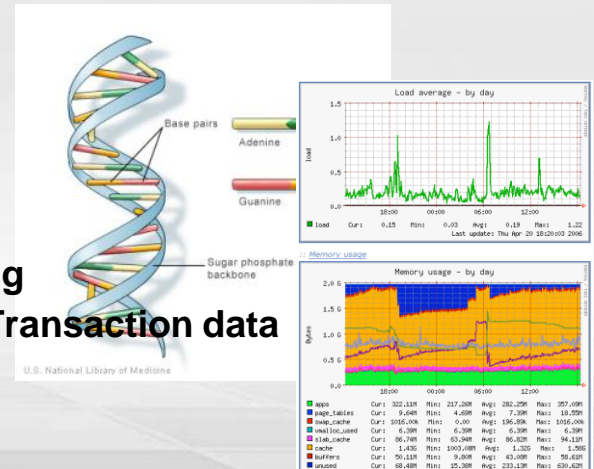
Open Source: Apache 2.0 Licensed



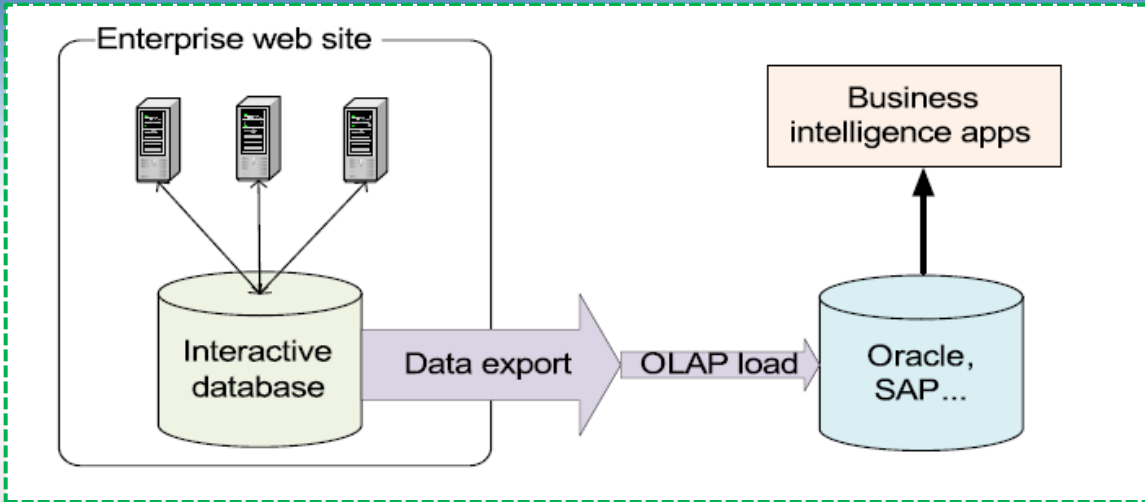
(MapReduce/ Hbase/ HDFS)

대용량 데이터 적용 분야

- Simulations and Scientific/Experimental Data, Medical imaging
- Existing Databases: Product catalogs, Historical sales data, Transaction data
- User Data: Web logs, Clicks on website, Pictures, Videos, etc
- Systems reporting status every second



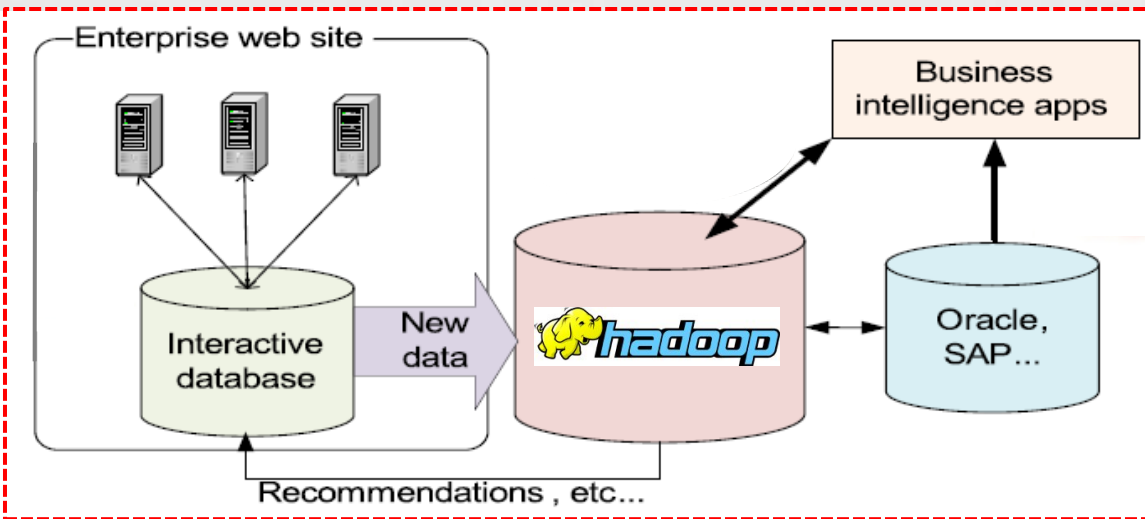
Cloud Computing 핵심 기술 - Hadoop (2)



RDBMS



- Total Data Volume : 100's TB
- High End Servers
- Licensing + Annual Maintenance



Hadoop

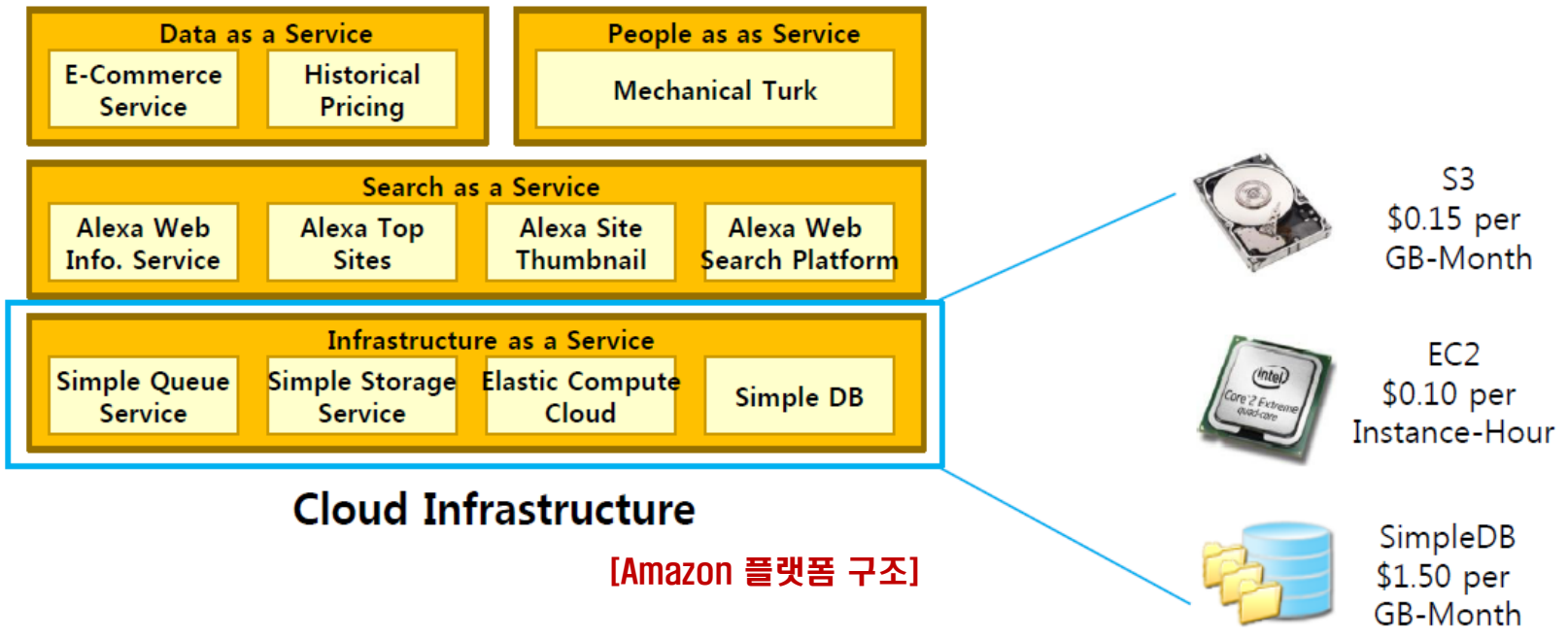


- Total Data Volume : 10's PB
- Commodity/ Utility Hardware
- Free Software + Support

(Source : <http://www.cloudera.com>)

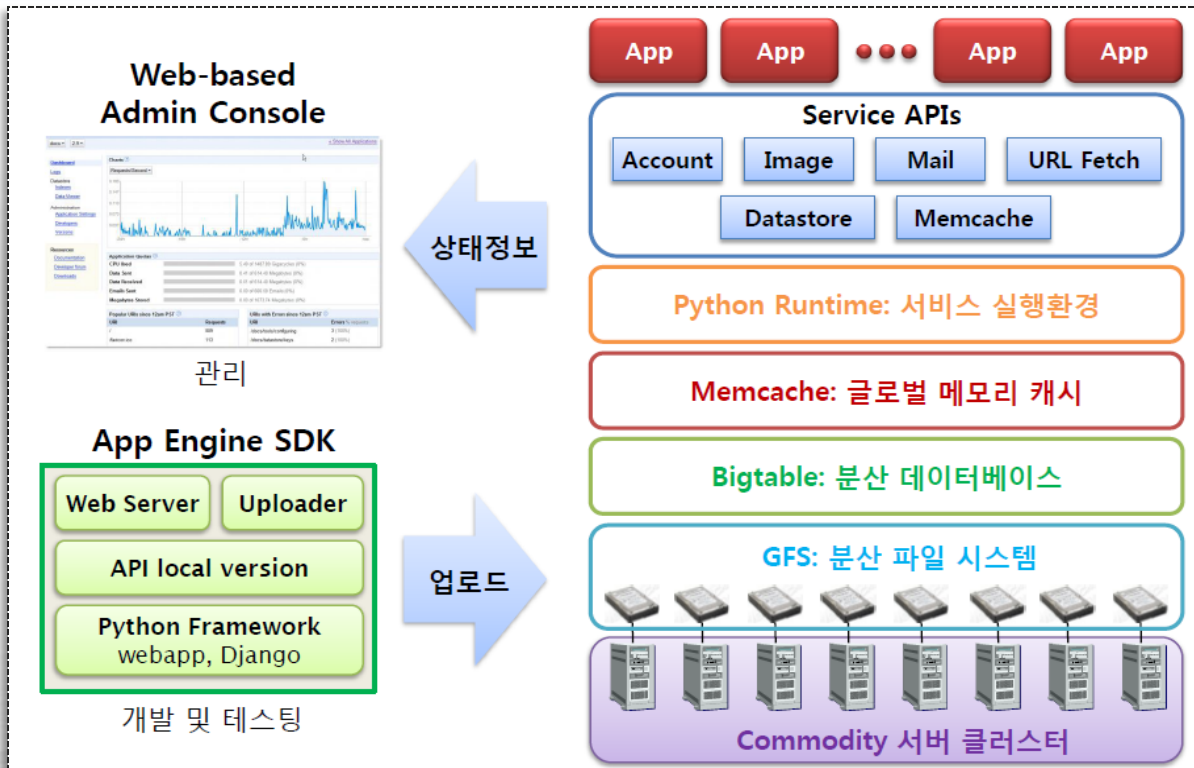
Amazon EC2 Platform

- 어플리케이션 컴포넌트와 플랫폼, 가상인프라 서비스 제공
- 2002년부터 AWS(Amazon Web Service)를 제공하면서 초기 클라우드 서비스 제공
- 2006년부터 유틸리티 컴퓨팅 개념을 도입한 EC2 (Elastic Compute Cloud)와 S3 (Simple Storage Service) 서비스 제공



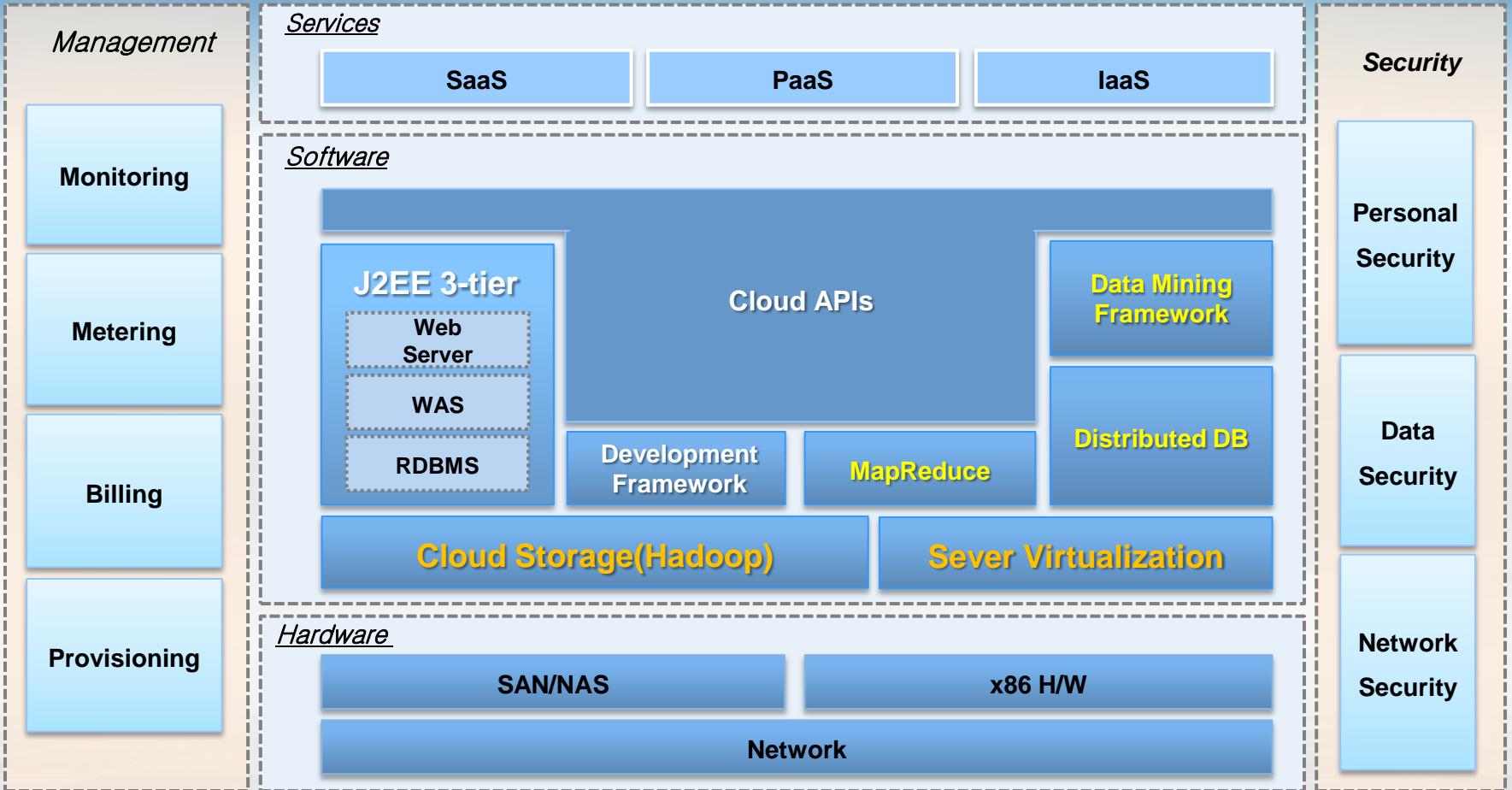
Google App Engine

- 웹기반 서비스 시장에서 SW 서비스 시장으로 진출 하였으며 점차 IT인프라 시장으로 확장
- 2008년에 선보인 App Engine은 구글 플랫폼 상에서 웹 어플리케이션을 자유롭게 개발할 수 있도록 하는 **PaaS (Platform as a Service)**를 제공



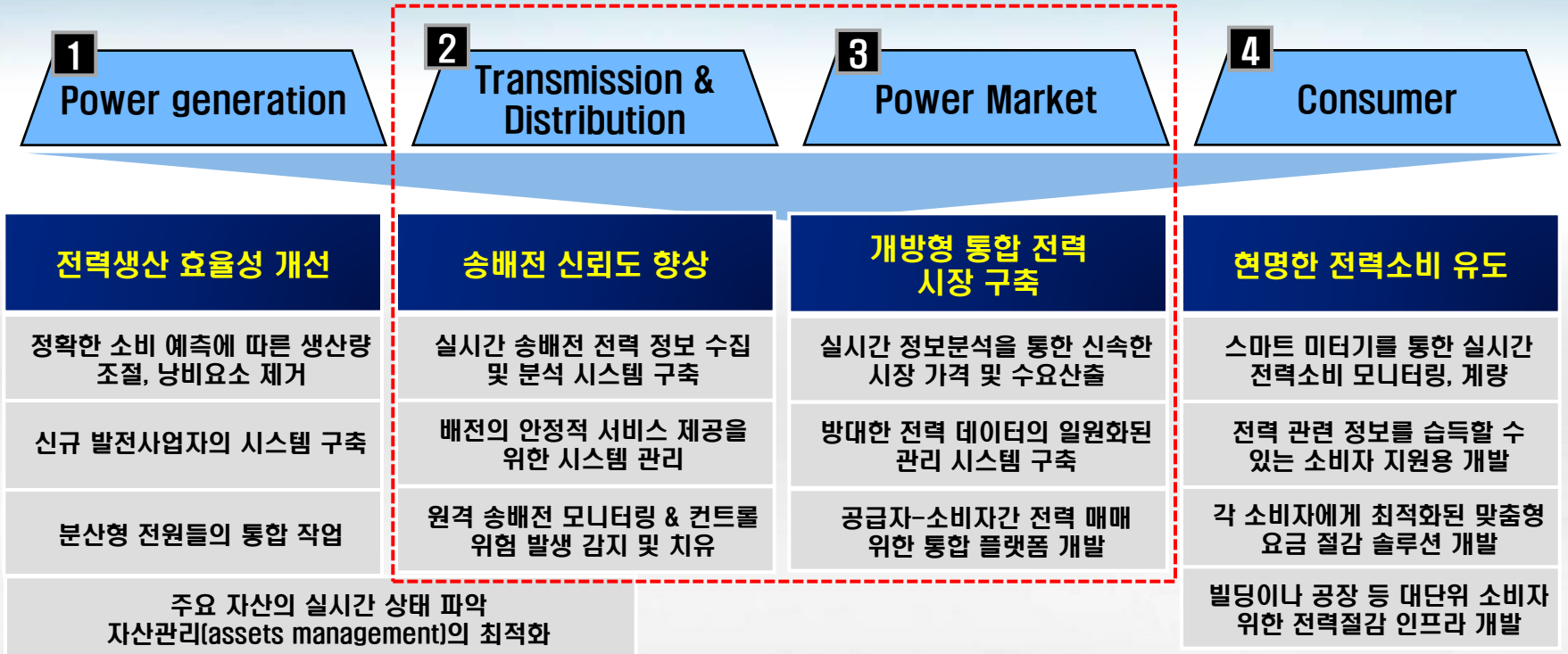
[Google App Engine 플랫폼 구조]

Cloud Computing 아키텍처 (SK C&C)

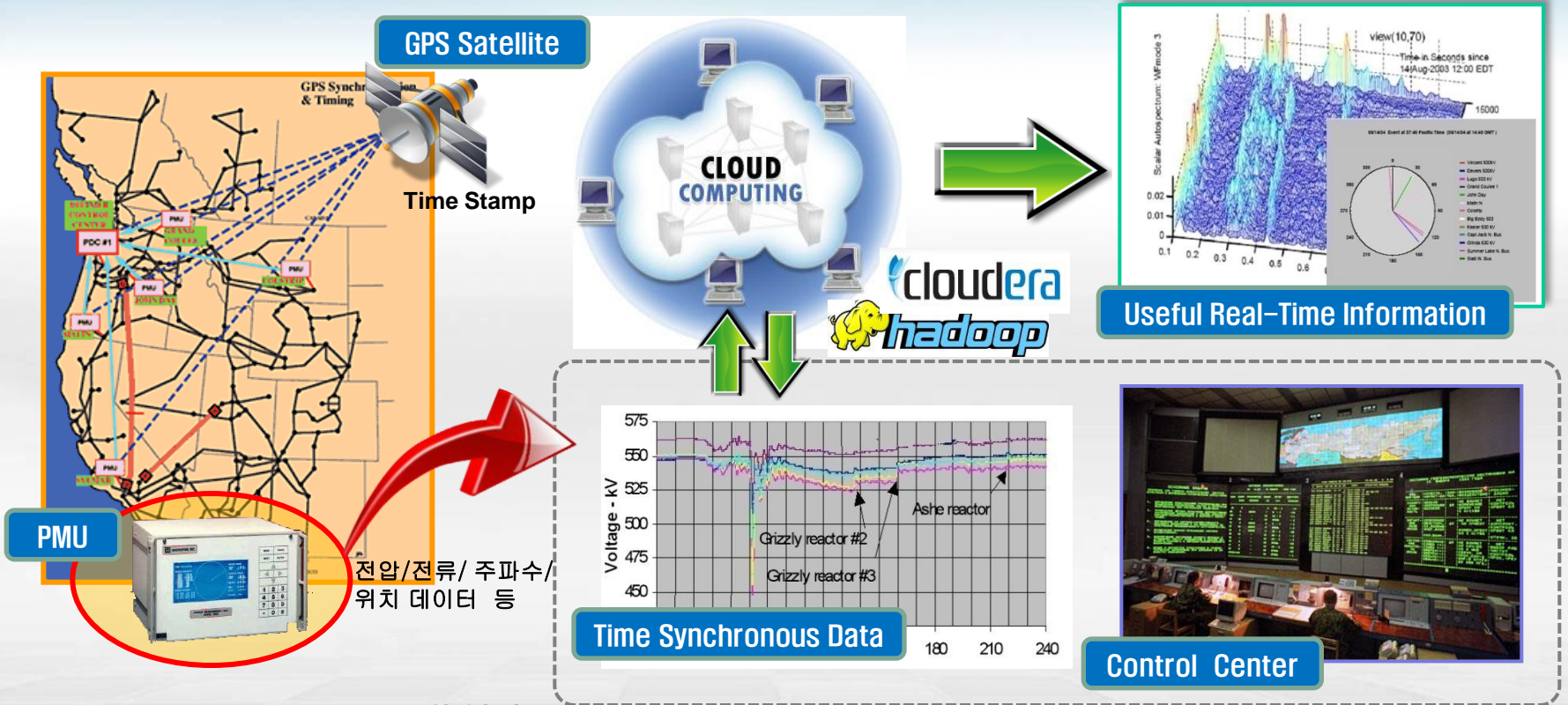


스마트 그리드 데이터 분석 플랫폼

스마트 그리드 분야 IT 개발 영역

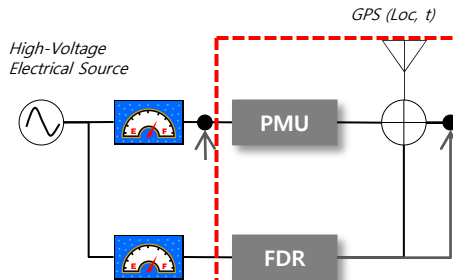


- ▶ 현재 미국 TVA(Tennessee Valley Authority)와 NERC(North American Electric Reliability Corp.)에서는 스마트 그리드의 발전, 송전, 배전 등 전력망 상태에 관한 데이터 분석, 처리하기 위해 클라우드 컴퓨팅 서비스 활용
- ▶ PMU 장치를 통해 수집한 전력망 데이터 분석을 위해 저렴한 컴퓨팅 파워를 이용해 대용량 데이터를 쉽고 안정적으로 처리할 수 있는 Hadoop의 분산 병렬 처리 기능 활용



PMU (Phasor Measurement Units, 위상측정기) 전력계통의 'Health Meter'로 불리는 PMU는 전력계통에 대해 MRI 수준의 정보를 제공하는 것으로 주어진 위치에서 초당 전압, 전류, 주파수 등의 정보를 측정함.

Power Generator & Transmitter



PMU data

- V (mag, angle): A(+), B(-), C(0) phases
- I (mag, angle): A(+), B(-), C(0) phases
- Freq.
- dF/dt
- Digitals: 1 or 2
- Status Flags

FDR data

Disturbance Info (floats)

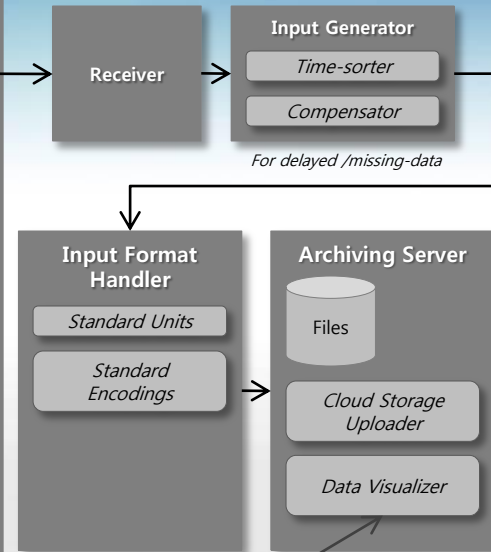
GPS data

- Timestamp (+Location)

Digital Data

- PMU : Phasor Measurement Unit
- FDR : Freq. Disturbance Recorder

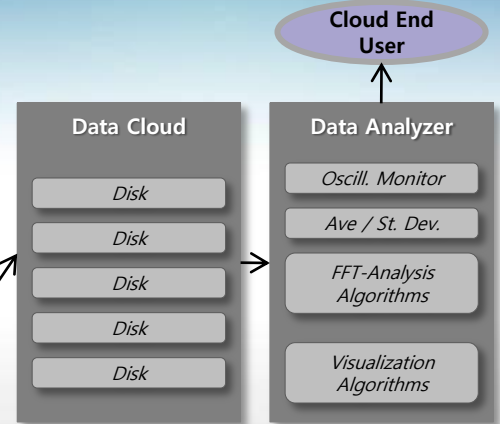
Power Data Collection & Check



End User

Standard Phasor data Protocol (IEEE C37.118-2005)

Power Data Storage & Processing



- FFT Analyses Algorithms
 - : Noise Reduction Analysis
 - : Real Signal Detection Analysis
- Power-grid Anomalies Detection Algorithms
- Rendering for Visualization Algorithms

표준

개발
기술

- 전력 계측기를 통한 실시간 전력 데이터 수집 및 전송
- GPS 데이터(시간 및 위치 정보) 활용 기술

기존 전력 데이터에 대한 변환 및 전송

- 실시간 대용량 데이터 수집 기술 (가용성 보장) 및 데이터 보정 기술
- 인코딩 및 Cloud Data 저장 기술 (다양한 프로토콜 지원)
- 모니터링/관리용 데이터 시각화 기술

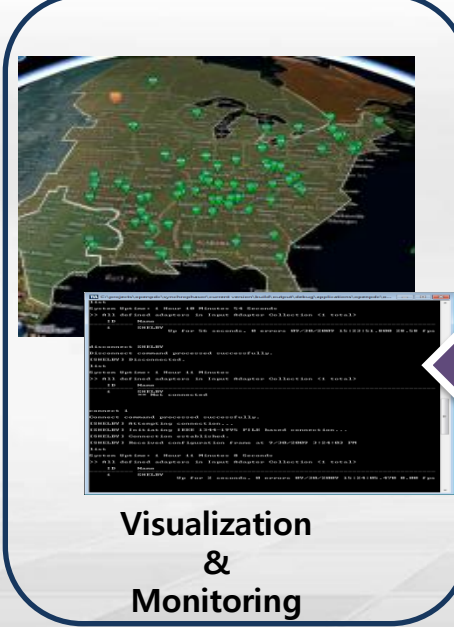
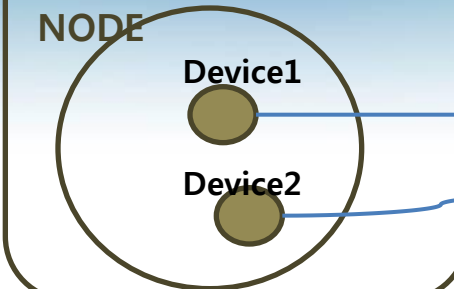
대용량 전력 데이터의 안정적인 수집

- 대용량 데이터 저장 기술 (Petabyte급 규모 확장성 고려)
- 데이터 안정성 확보 위한 다중 복제기술
- 분산병렬처리기술 및 Data 분석 알고리즘

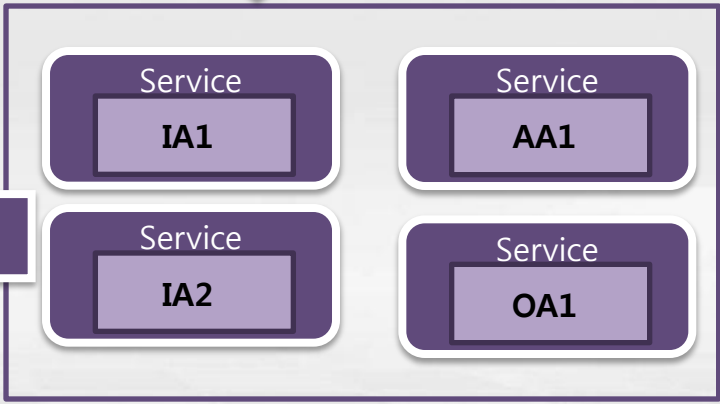
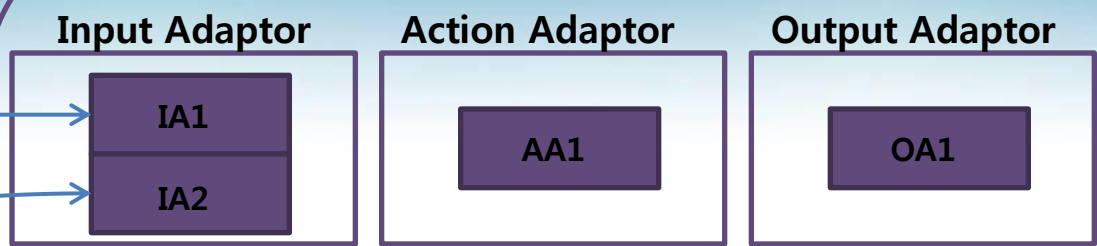
분산병렬컴퓨팅 기술 통한 데이터 분석

OpenPDC Collector 개요

Physical environment

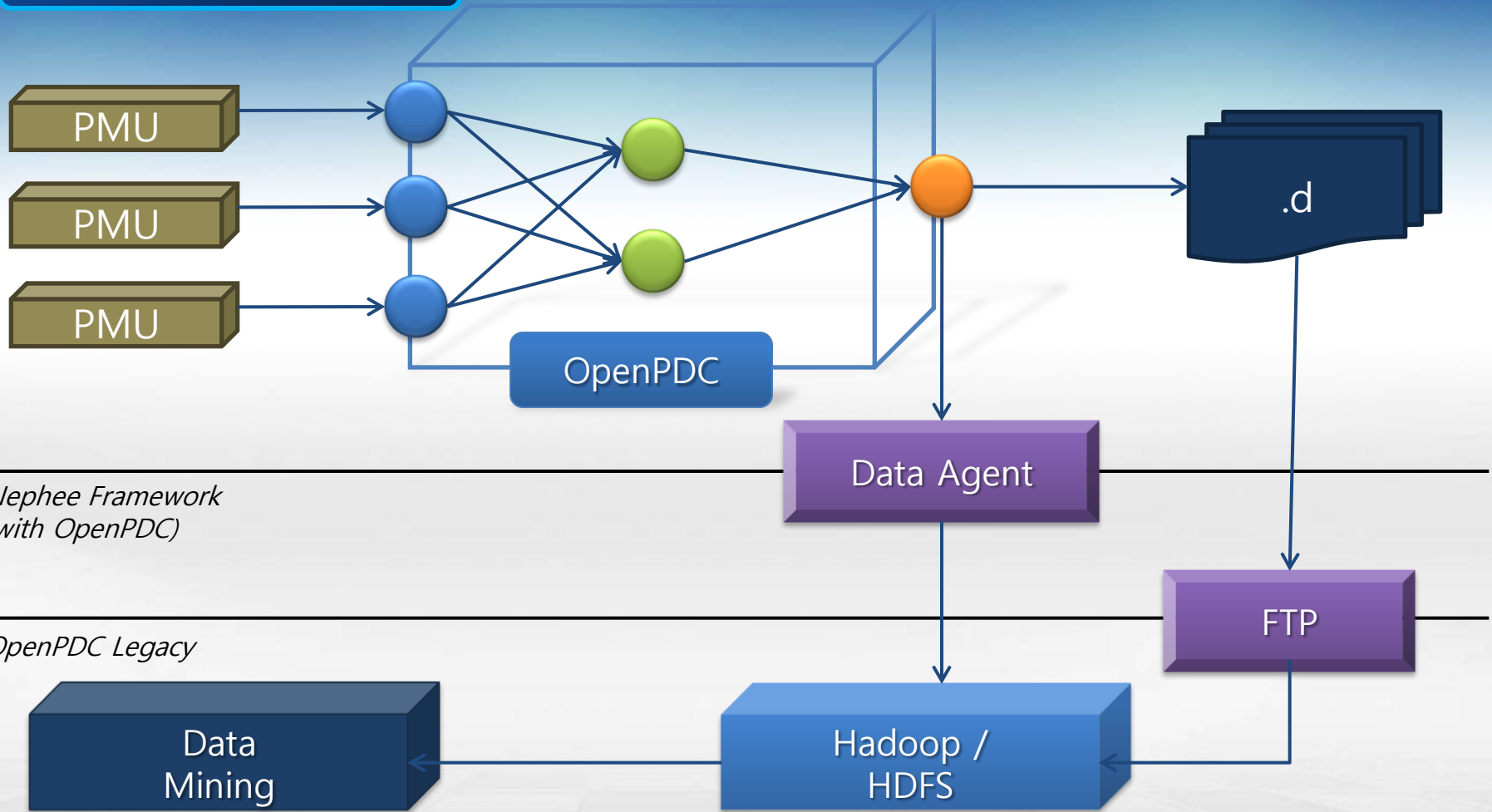


Logical Environment



OpenPDC

OpenPDC Architecture



스마트그리드 데이터 분석 플랫폼
Architecture (Nephee 프로젝트)

