

클라우드 웹서비스 개발자를 위한 Boot Camp

SW공학센터 SW공학기술팀

장원석 책임 연구원

wsjang@nipa.kr

2011.4.27

시간	내용
13:30~14:00	등록
14:00~15:00	클라우드 컴퓨팅 개요 및 기술 - 클라우드 컴퓨팅 개념 및 사례 - 클라우드 컴퓨팅 기술 소개
15:00~15:30	Break Time
15:30~18:00	클라우드 컴퓨팅 개발 실습 - 클라우드 서비스 체험 실습 - AWS API 동작 실습

1. 클라우드 컴퓨팅

- 클라우드 컴퓨팅 개요
- 클라우드 솔루션 소개

2. 클라우드 기술 소개

- 기술 1. NoSQL
- 기술 2. 파일시스템 I/O 가상화
- 기술 3. 분산 관리 (Chubby)
- 기술 4. 병렬 처리 MapReduce
- 기술 5. 로그 관리 Scribe
- 기술 6. 캐시 (Memcached)

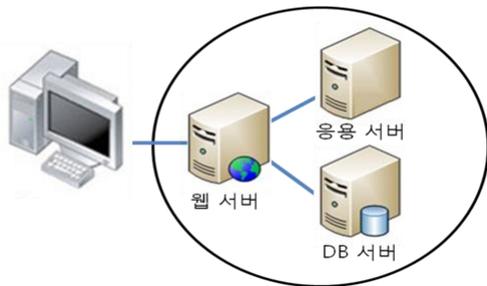
3. 클라우드 컴퓨팅 실습

- 클라우드 서비스 체험
- AWS API 동작 실습
- Q&A

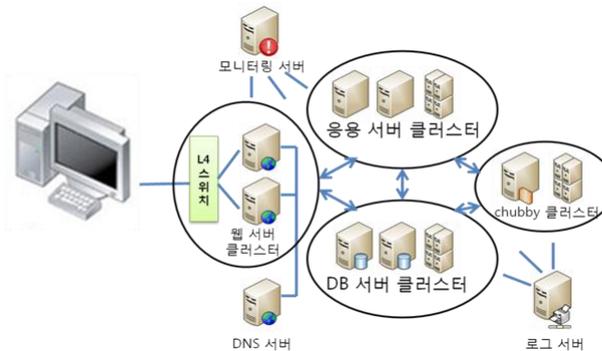
1. 클라우드 컴퓨팅이란?

- 클라우드 컴퓨팅은 인터넷 기술을 이용하여 내외부 고객들에게 확장성 있고 탄력적인 IT 서비스가 제공되는 방식이다. (Cloud is a style of computing where scalable and elastic IT-related capabilities are provided as a service to customers using Internet technologies.)
- *by Gartner Report*

[서비스 컴퓨팅 환경의 진화]



[단순 웹 환경]



[클러스터 환경]



[클라우드 환경]

- 클러스터 컴퓨팅 : 여러 대의 컴퓨터를 하나로 묶음
- 그리드 컴퓨팅 : 분산 병렬 처리를 가능하게 함
- 클라우드 컴퓨팅 : 그리드 컴퓨팅을 간편하게 구현하기 위한 프레임워크를 갖춘

<< 특징 >>

1. 확장성

- 서비스 규모를 얼마든지 확장할 수 있음

2. 탄력성

- 서비스의 확장/축소/복구에 시간/비용이 적게 소요됨

3. 애매모호성

- 어느 서버에서 어떤 서비스가 실행되는지 쉽게 알 수 없음

1. 클라우드화 6단계

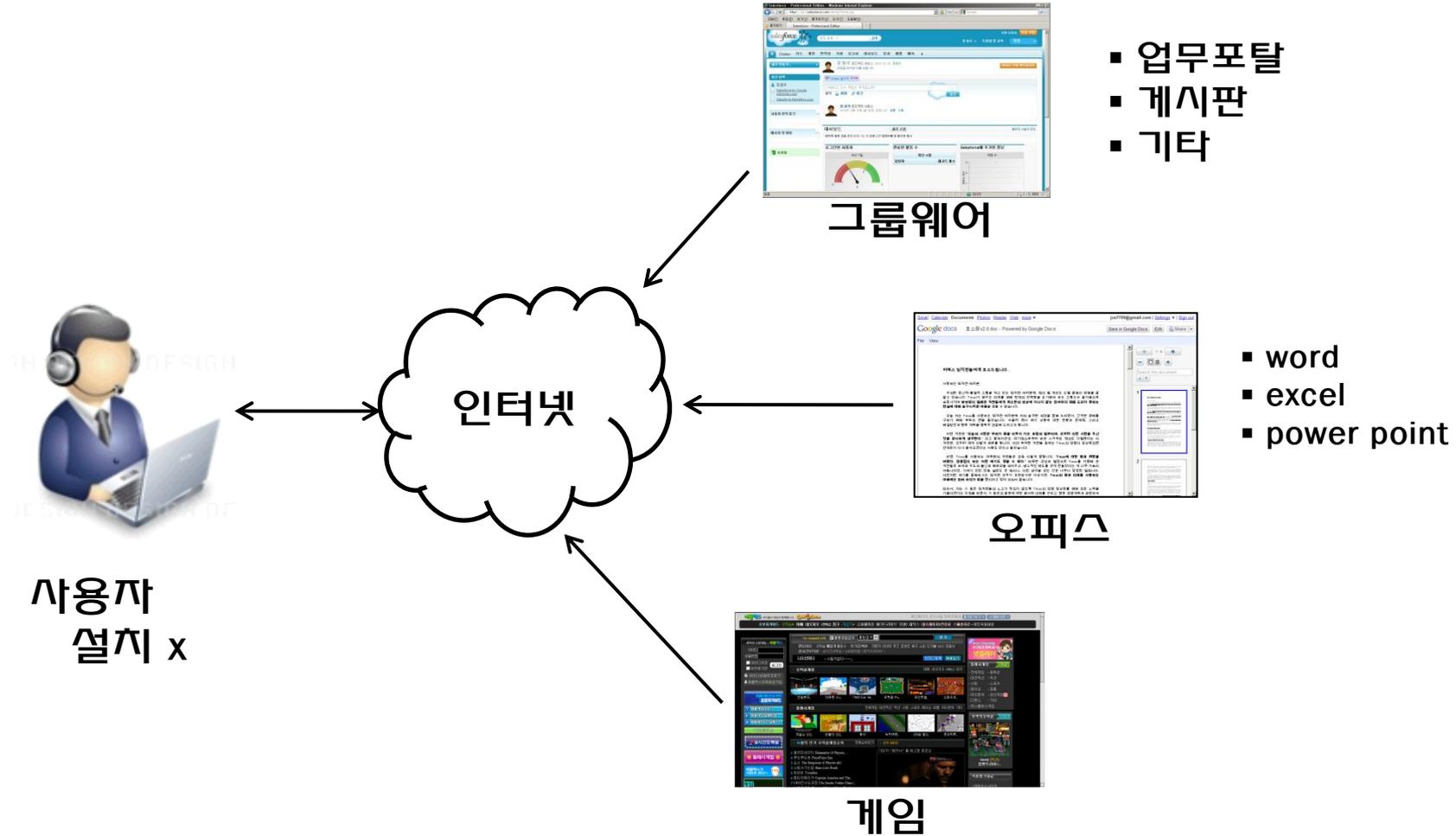


**클라우드 컴퓨팅은
기존 환경의 문제점 해결을 위해
보다 성숙한 구조를 지향**

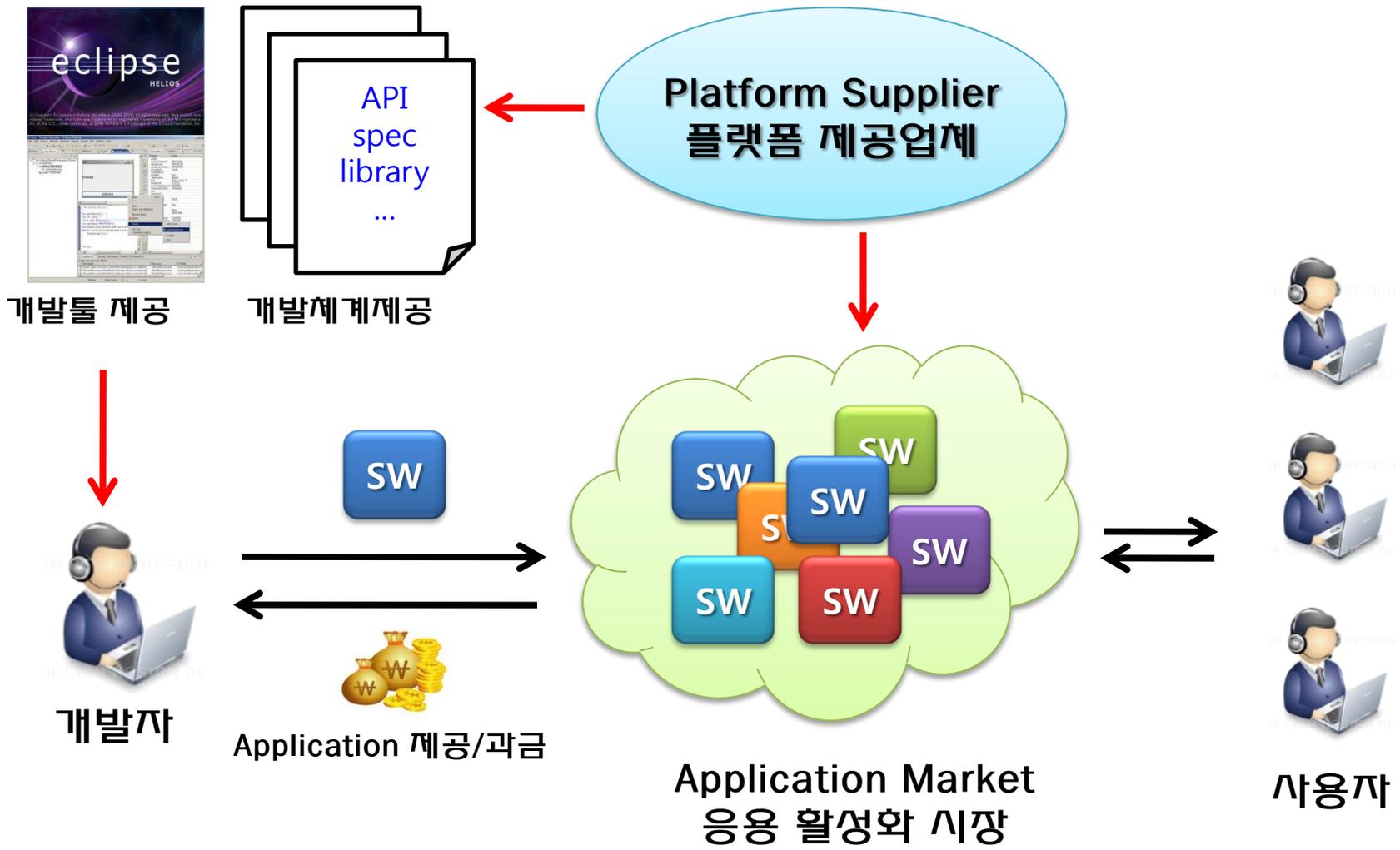
1. SaaS / PaaS / IaaS



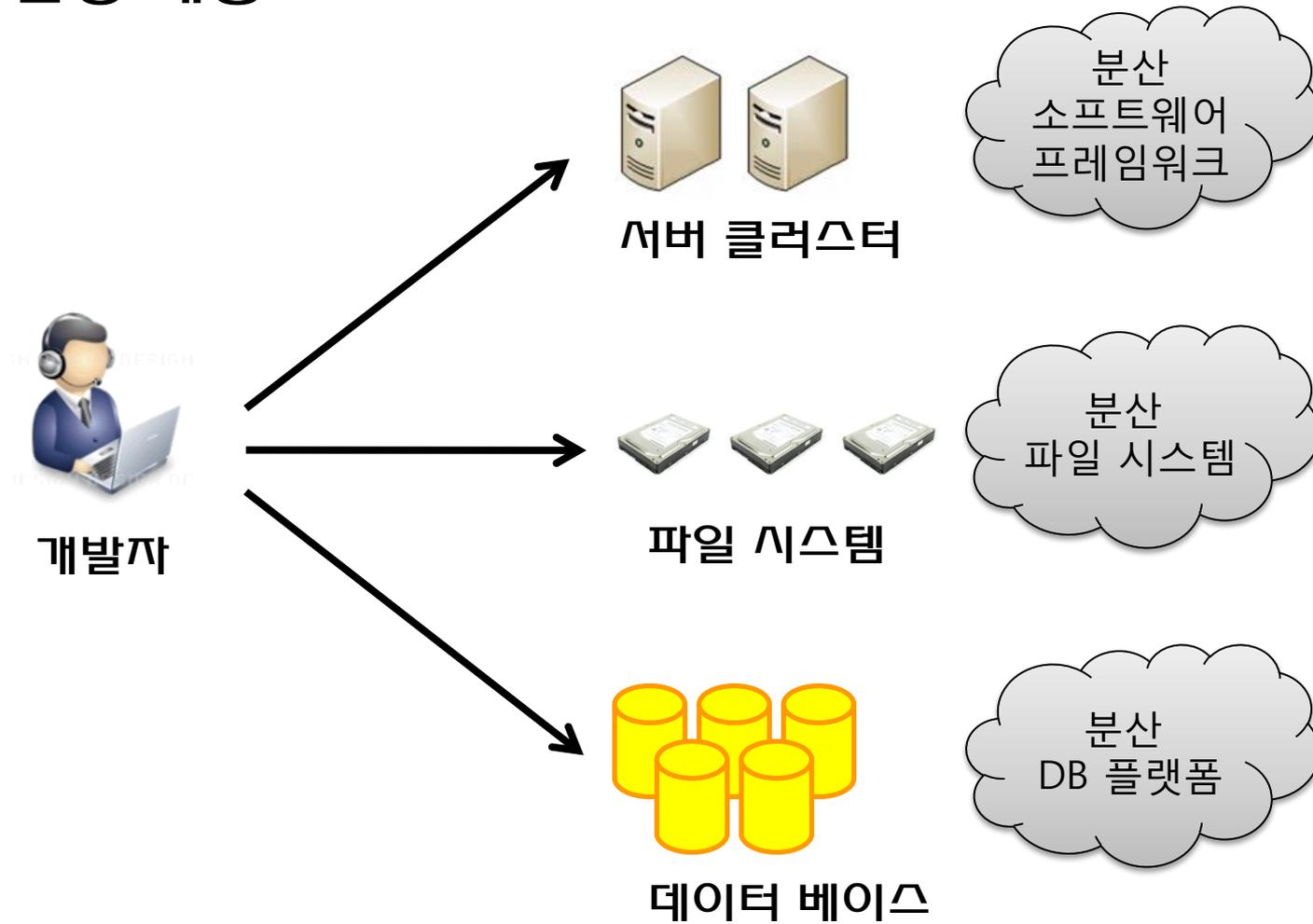
1. 인터넷의 서비스를 빌리는 방식



1. 개발 플랫폼 제공을 통한 자사 솔루션 시장 유치



1. 분산 환경 제공



1. Private Cloud

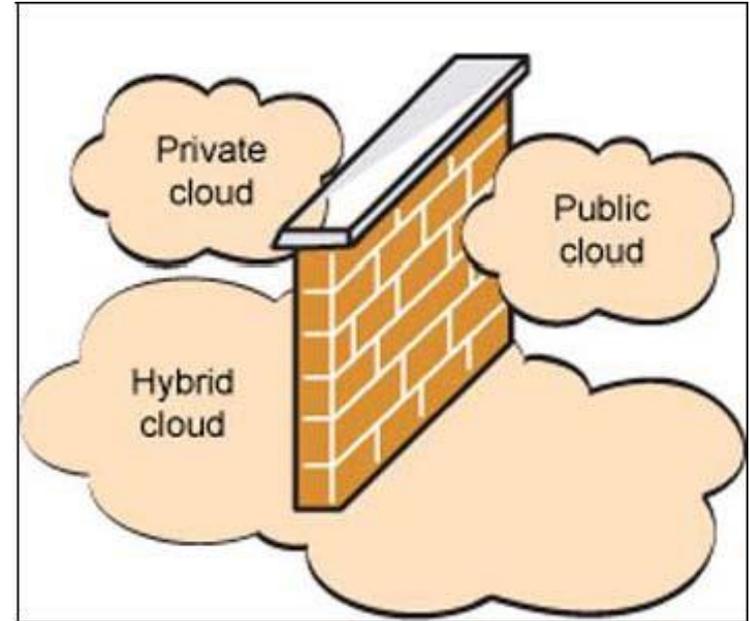
- 내부에만 서비스

2. Public Cloud

- 외부에도 서비스

3. Hybrid Cloud

- private 와 public 을 합쳐서 서비스



1. 클라우드 컴퓨팅

- 클라우드 컴퓨팅 개요
- 클라우드 솔루션 소개

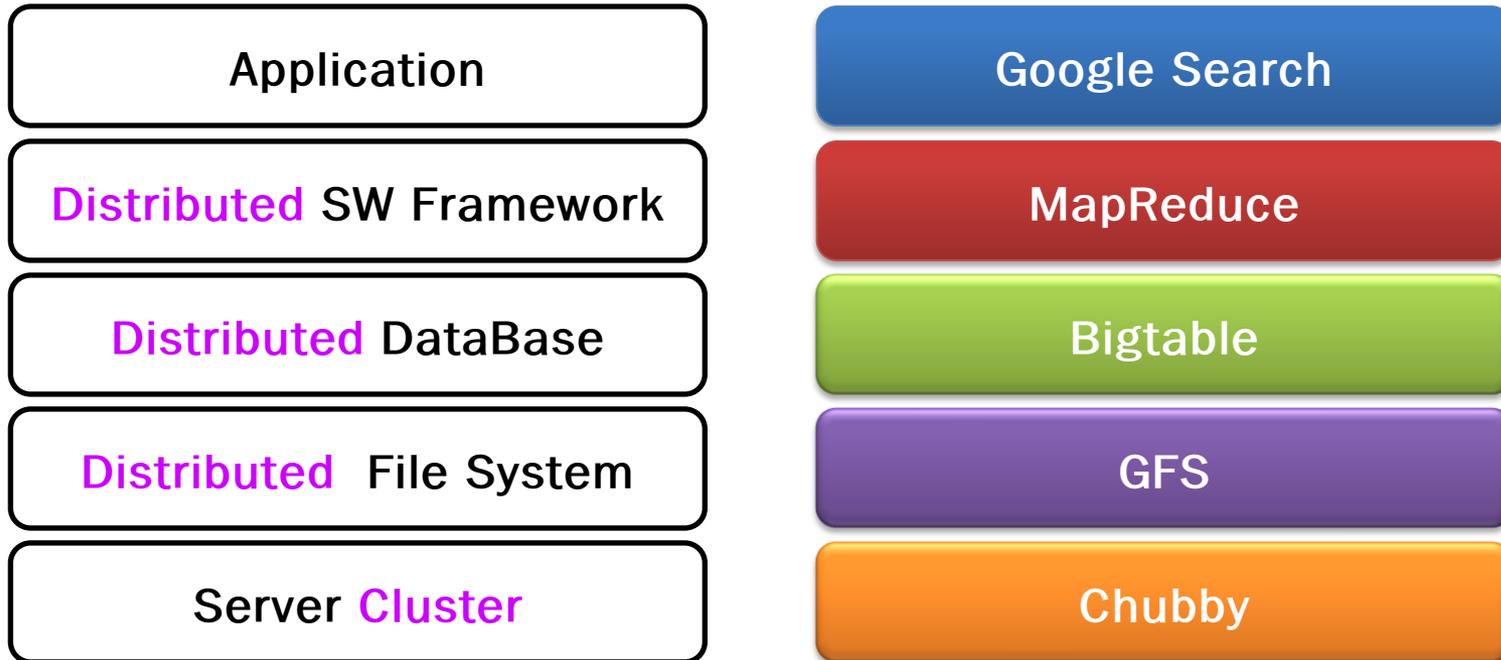
2. 클라우드 기술 소개

- 기술 1. NoSQL
- 기술 2. 파일시스템 I/O 가상화
- 기술 3. 분산 관리 (Chubby)
- 기술 4. 병렬 처리 MapReduce
- 기술 5. 로그 관리 Scribe
- 기술 6. 캐시 (Memcached)

3. 클라우드 컴퓨팅 실습

- 클라우드 서비스 체험
- AWS API 동작 실습
- Q&A

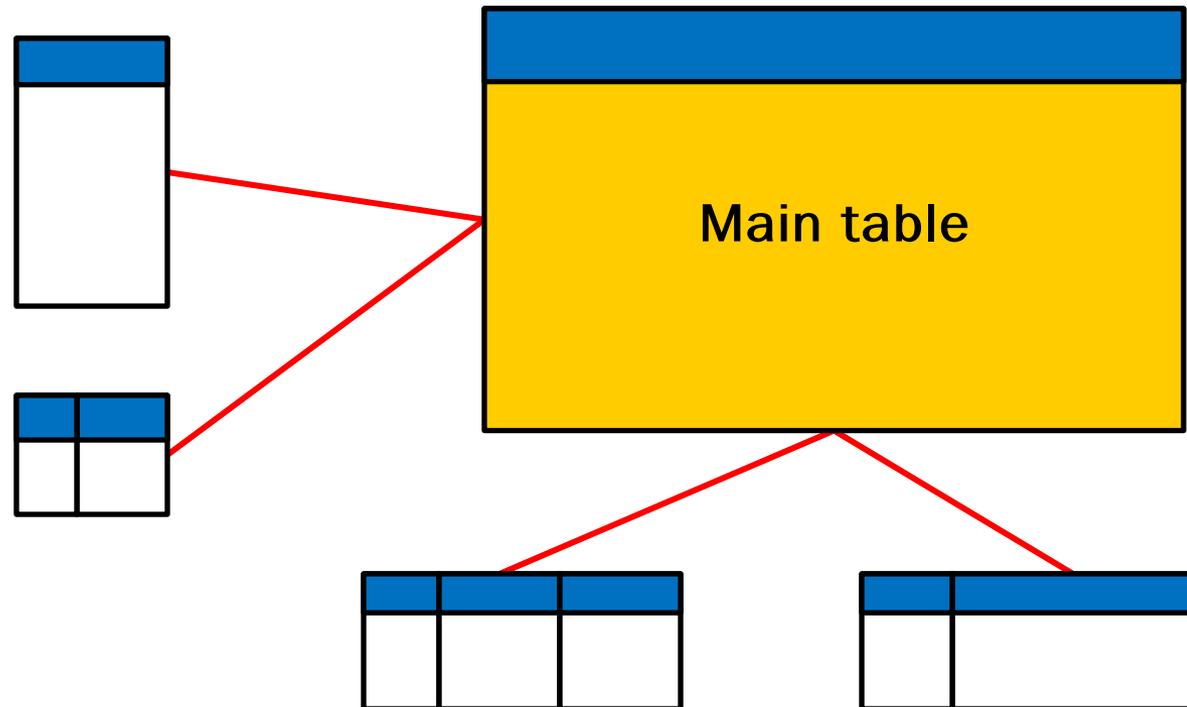
1. Google 에서 발표한 분산 플랫폼 기술



- ✓ 구글이 저가의 비용으로 고용량의 검색 솔루션을 개발/운영/관리할 수 있게 했던 원동력이 되는 기술
- ✓ 논문 공개를 통하여 오픈소스화 (Hadoop) 되어 널리 사용됨
- ✓ NoSQL (Bigtable) , 파일 I/O 가상화 (GFS) , 분산관리 (Chubby) , 병렬처리 (MapReduce) 등 구글 기술 기반

1. 관계형 데이터베이스 모델

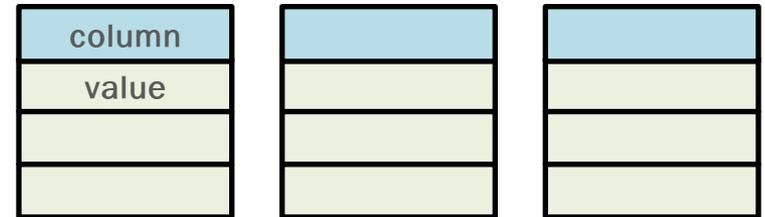
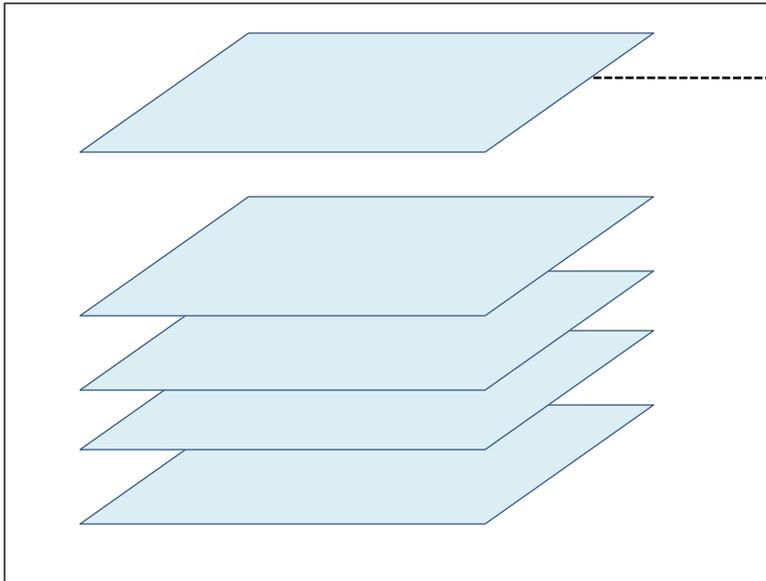
- 관계형 데이터 베이스 (Oracle, MS-SQL, ...)
 - 데이터 정합성 증시
 - 데이터 관계 증시



1. 키-값 데이터 모델

- 분산 데이터 베이스 (H-base, Big table, 클라우드데이터, ...)

table



(좌표 → 데이터) 모델
(row , column) = value

관계형 데이터베이스

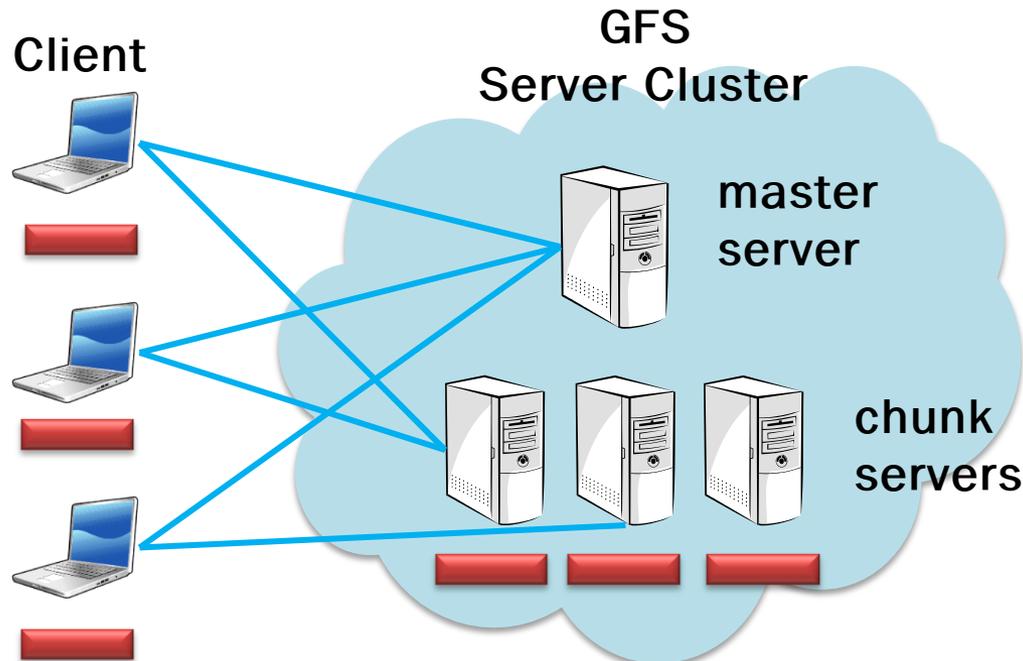
```
select * from table  
where table.no=10
```

NoSQL 데이터베이스

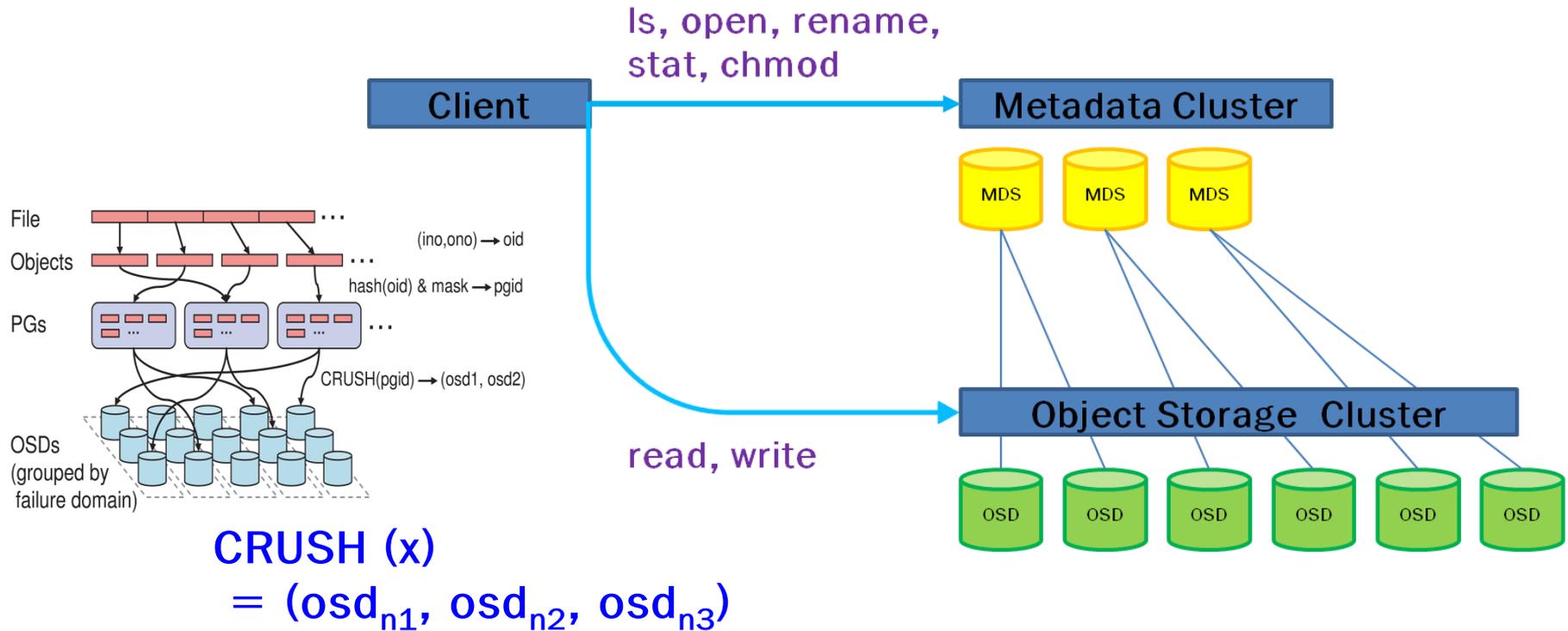
```
Table.get (10, column1 , column2 );
```

1. 분산 파일 시스템 (GFS)

- RAID 기술을 분산 시스템으로 구현
- 데이터 사본을 3개 유지
- 안정성 (Fault Tolerance)과 성능 (High Performance) 중시
- 대용량 확장 가능 파일 시스템



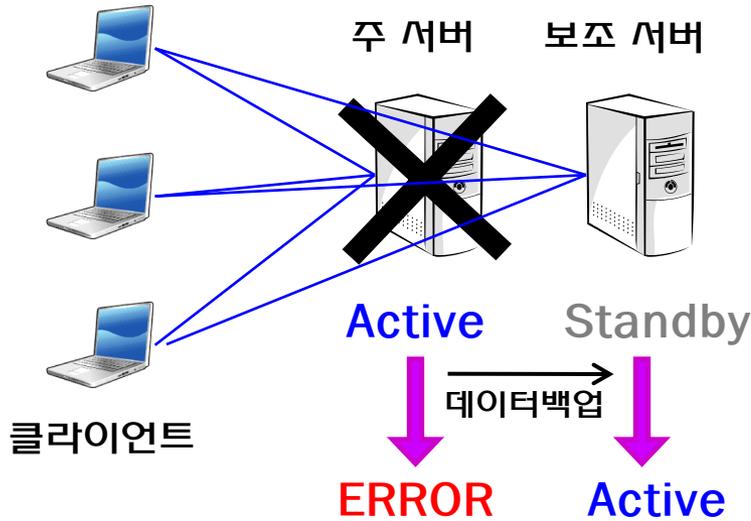
1. 분산 파일 시스템 구조 (CEPH)



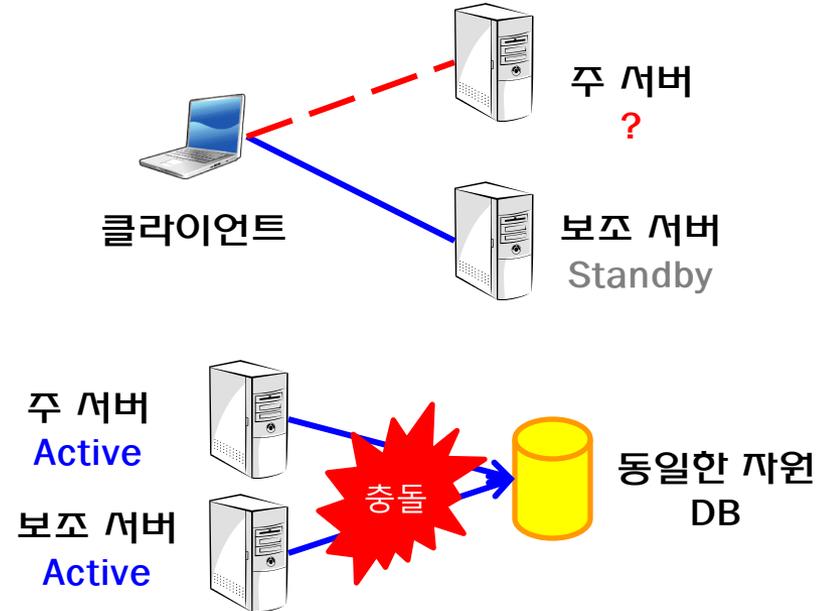
- ✓ OS kernel file system 이 수정됨
- ✓ file 을 object 단위로 분할하여 object id 로 관리

1. 단순 분산 관리

이중화 서버 구조

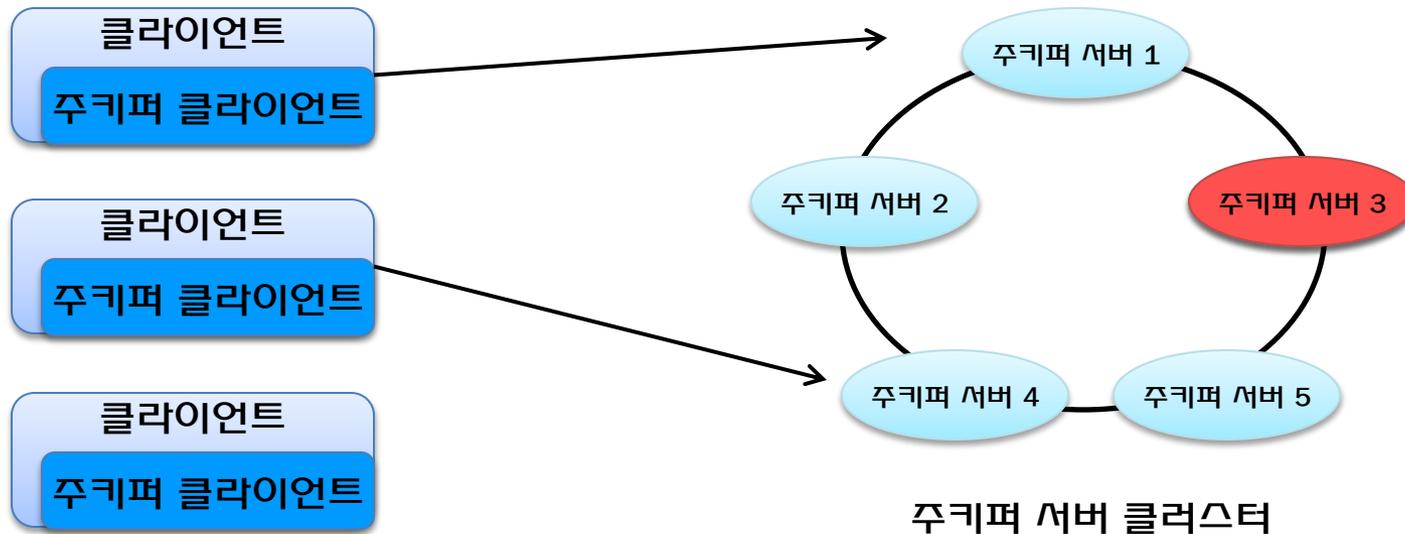


문제점



- ✓ 보다 체계적이고 확장 가능한 분산 관리 구조가 필요
- ✓ 어플리케이션 차원에서 복잡하게 해결하기 보다는 시스템 아키텍처로서 문제점을 극복

1. ZooKeeper 구조 (Chubby 분산락관리)

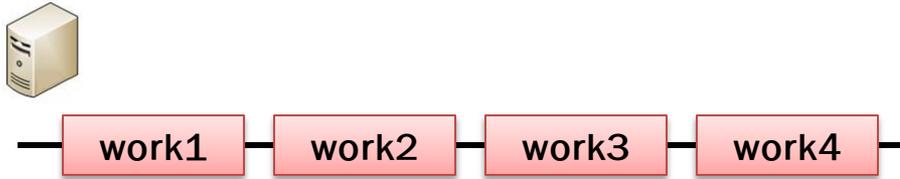


- ✓ 주키퍼 서버 클러스터는 항상 데이터를 동일하게 복제, 관리
- ✓ 주키퍼 서버 클러스터에서 리더(leader)를 투표를 통해 선출
- ✓ 리더는 다른 서버들의 동기화 및 락 관리의 기준이 됨

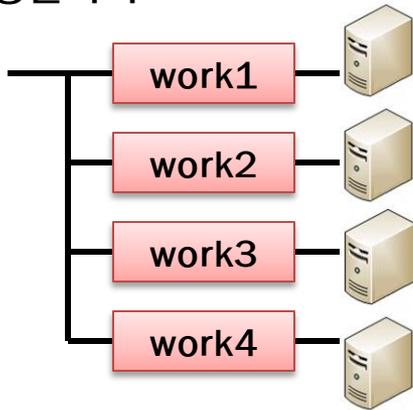
1. 병렬 처리 (Parallel Processing)

- 대용량, 복잡한 연산 처리
- 그리드 연산을 통해 슈퍼컴퓨터 성능을 구현

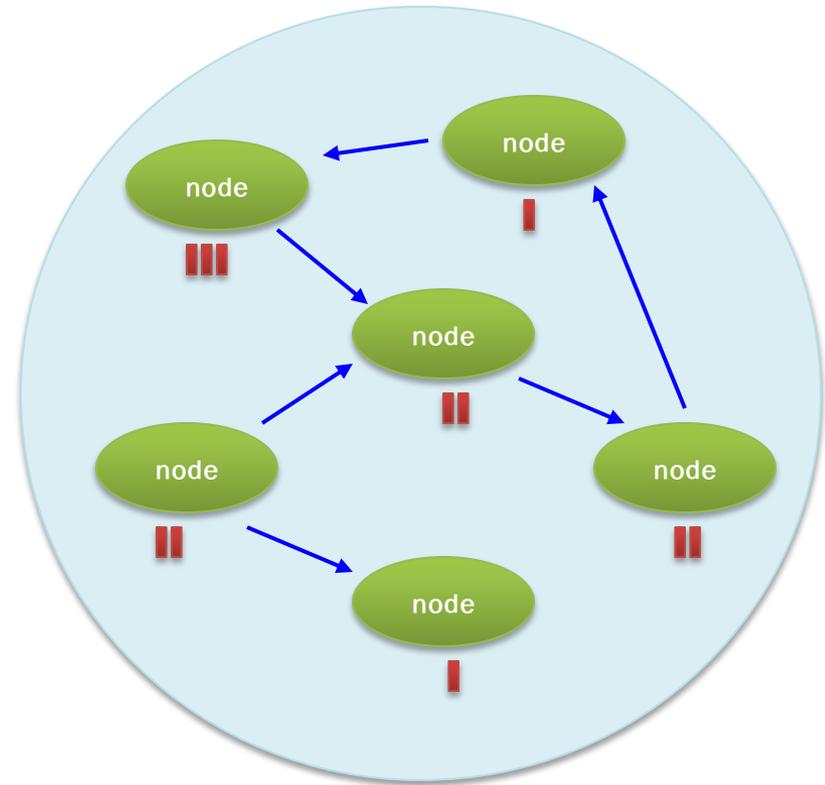
직렬처리



병렬처리

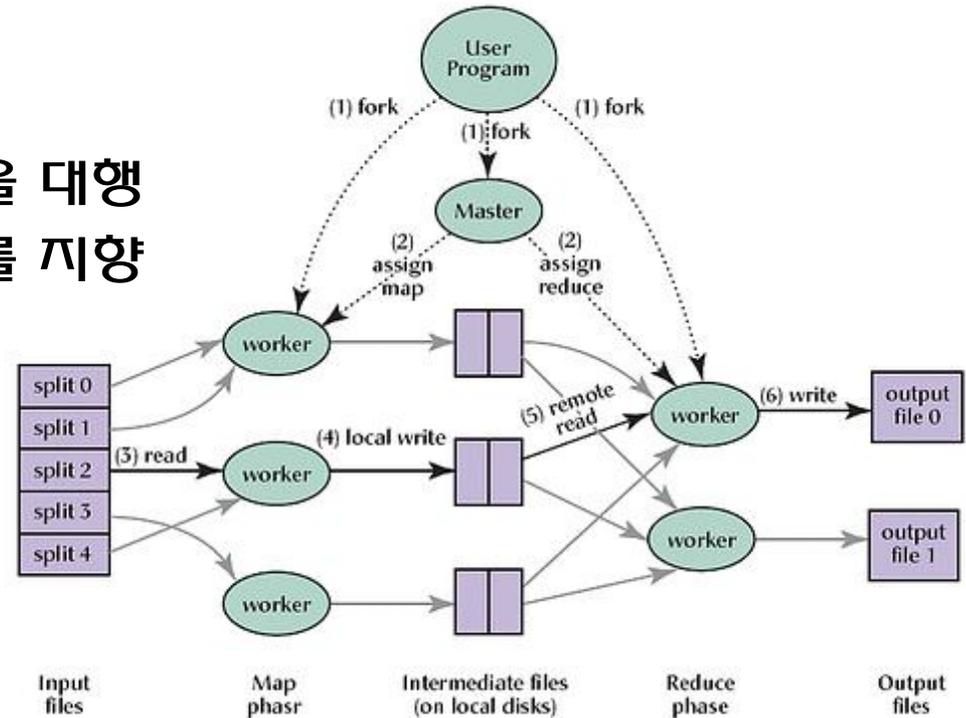


MPI 방식 (Message Passing Interface)



1. MapReduce

- 키-값 모델
- 데이터를 나누고, 합치는 작업을 대행
- 병렬 처리 중에서 간단한 구조를 지향



예시) SORT

899, 12, 34,
67, 1024, 512,
128, 1013,
2009, 666,
1225, 111

M

1	12 - 34 - 899
2	67 - 512 - 1024
3	128 - 1013 - 2009
4	111 - 666 - 1225

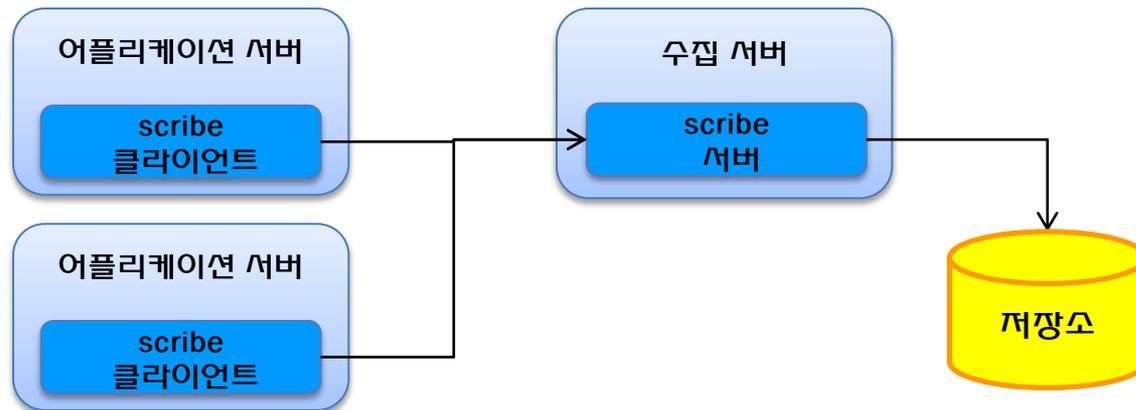
R

12 - 34 - 67 -
111 - 128 - 512
- 666 - 899 - ...

1. 로그의 수집과 분석

- 단순 어플리케이션 : 디버깅 활용, 장애 원인 파악
- 대규모 웹서비스 :
 - 어플리케이션 상태 정보
 - 사용자 이용 통계 내역 패턴 분석
 - 장애 예측, 자동 복구 파악

- 스크라이브 (Scribe)



- ✓ 부하가 적은 경량 어플리케이션 선호

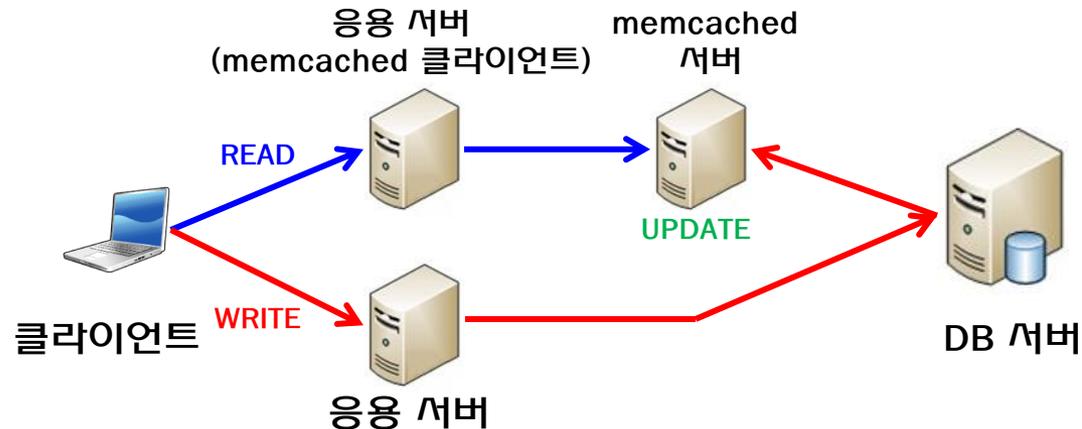
1. 캐시 (cache)

- 캐시의 사용 : 빠른 응답 속도



2. Memcached

- 데이터베이스 원격 분산 메모리 캐시 시스템
- 분산 시스템의 약점 : 속도
- 내부 구조 : 네트워크 해시 테이블



1. 클라우드 컴퓨팅

- 클라우드 컴퓨팅 개요
- 클라우드 솔루션 소개

2. 클라우드 기술 소개

- 기술 1. NoSQL
- 기술 2. 파일시스템 I/O 가상화
- 기술 3. 분산 관리 (Chubby)
- 기술 4. 병렬 처리 MapReduce
- 기술 5. 로그 관리 Scribe
- 기술 6. 캐시 (Memcached)

3. 클라우드 컴퓨팅 실습

- 클라우드 서비스 체험
- AWS API 동작 실습
- Q&A

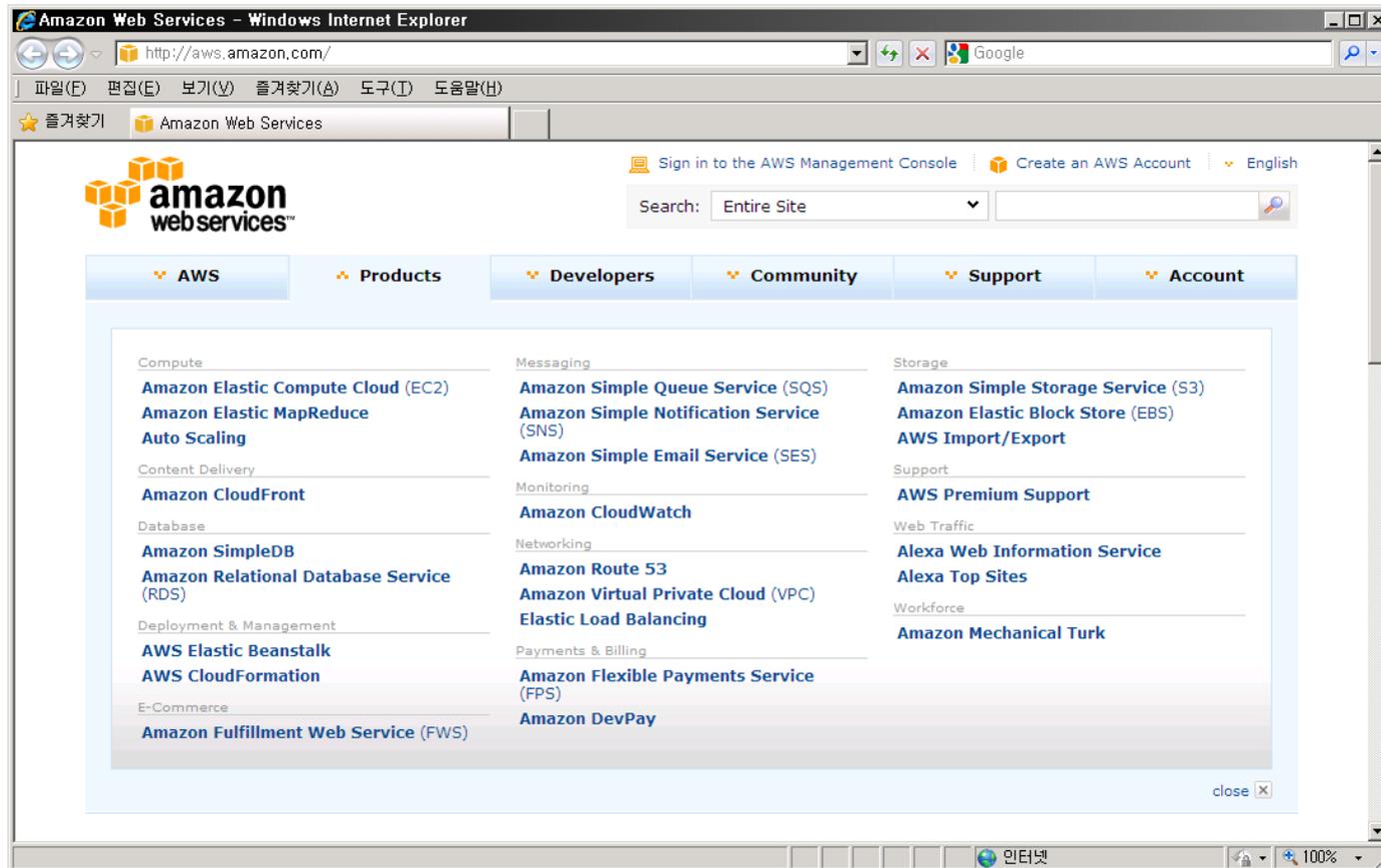
1. 세일즈 포스 닷컴

- 인터넷을 통해 그룹웨어 서비스를 제공 (SaaS)
- <http://www.salesforce.com>



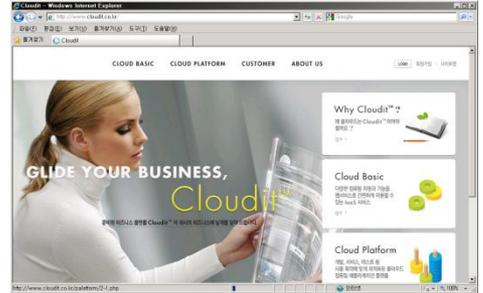
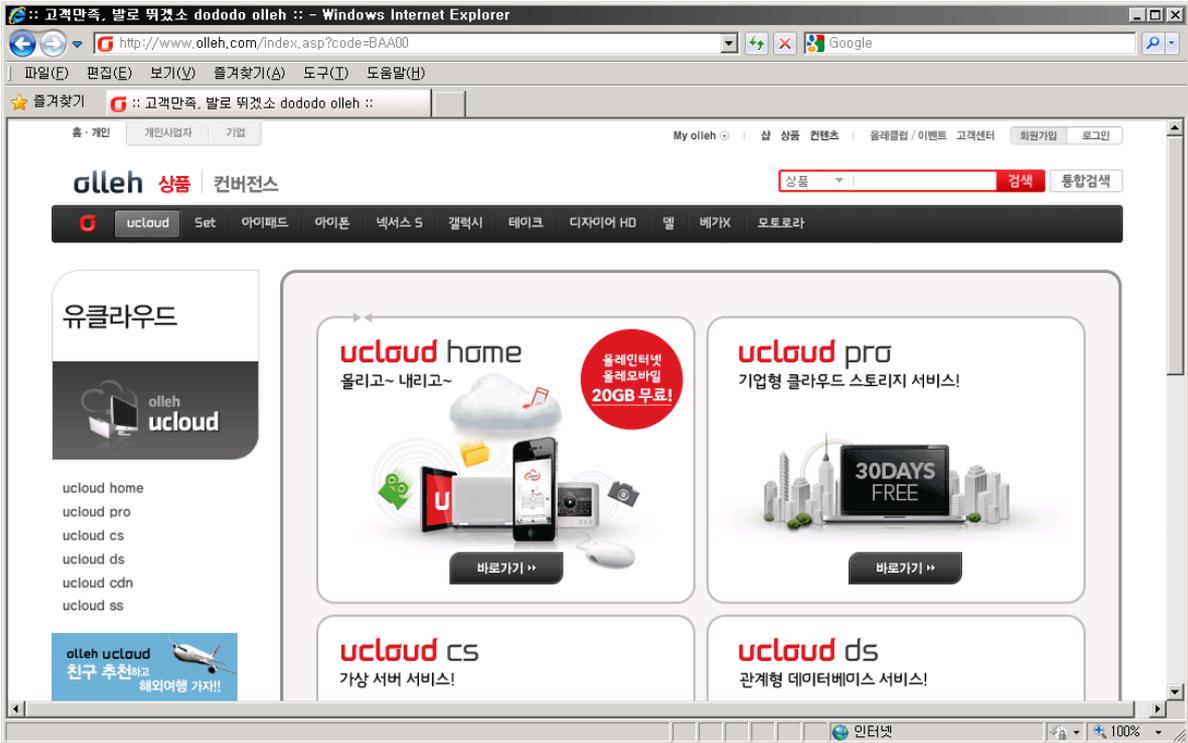
1. Amazon Web Service

- Database , Storage 등 여러가지 서비스를 제공



1. 국내서비스

- KT : 유클라우드 (<http://www.olleh.com/index.asp?code=BAA00>)
- 이노그리드 : CLOUDIT (<http://www.cloudit.co.kr/>)
- 호스트웨이 : FlexCloud (http://www.hostway.co.kr/cloud_flexcloud.html)



1. AWS API

- REST (Representational State Transfer) API
- SOAP (Simple Object Access Protocol) API
- REST 방식은 간결하지만 부수적인 (보안,정책) 구현이 필요, SOAP 방식은 XML로 플랫폼에 독립적이고 분산 환경에 적합하지만 복잡함
- REST 방식이 85% 압도적임
 - GET Service
 - **DELETE Bucket**
 - **GET Bucket**
 - List Multipart Uploads
 - **PUT Bucket**
 - **DELETE Object**
 - **GET Object**
 - HEAD Object
 - POST Object
 - **PUT Object**
 - Initiate Multipart Upload
 - Upload Part
 - Complete Multipart Upload
 - Abort Multipart Upload
 - List Parts
 - ...

1. PUT Object

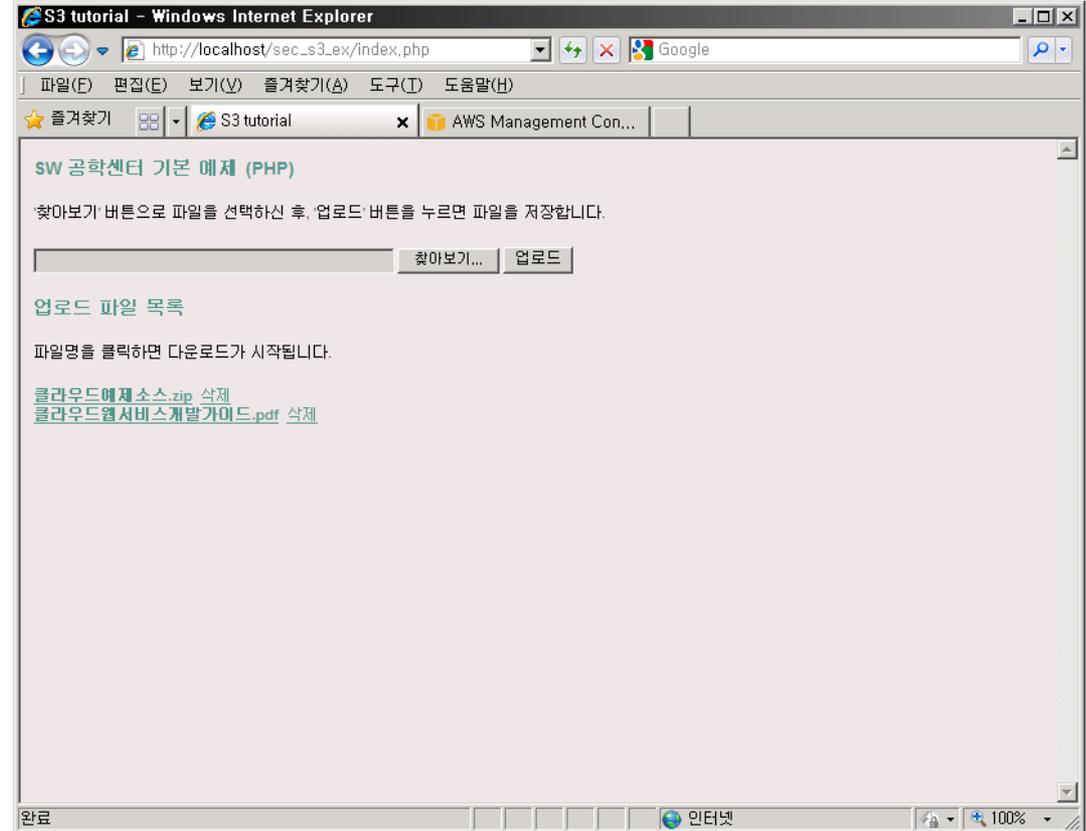
- 파일 업로드

2. GET Object

- 파일 다운로드

3. DELETE Object

- 파일 삭제

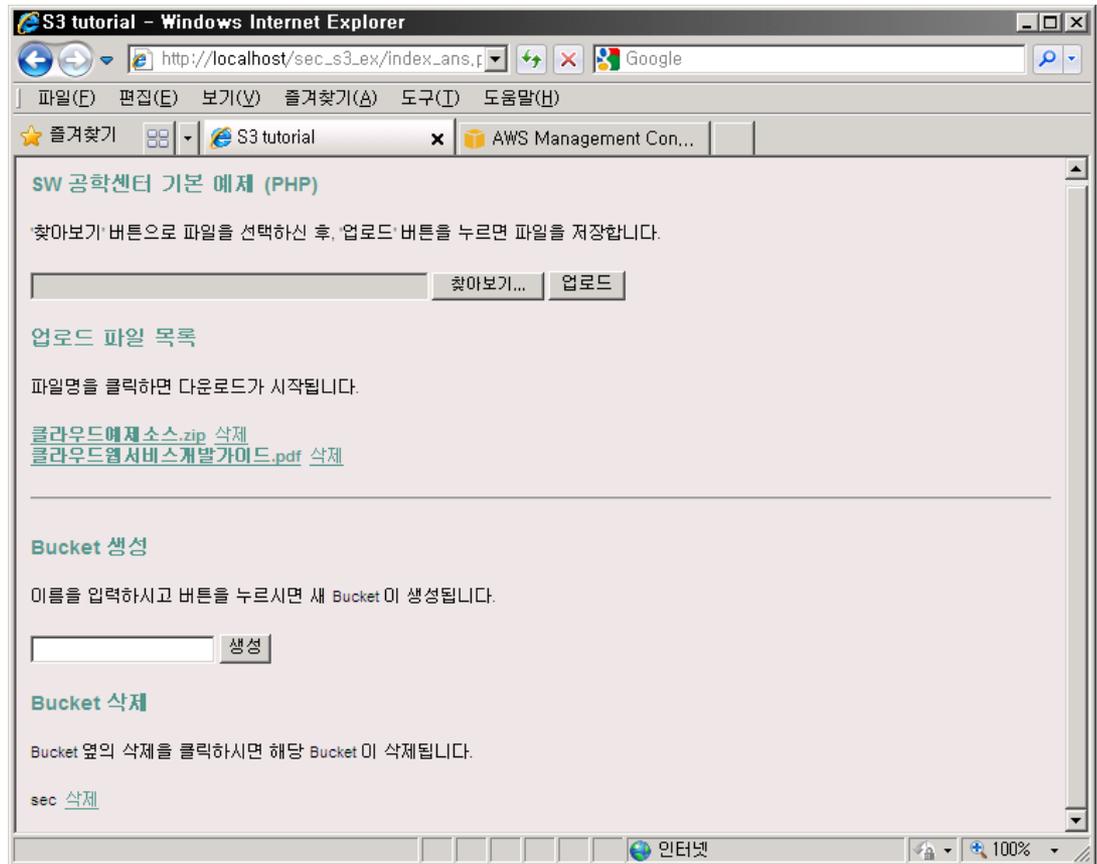


1. PUT Bucket

– Bucket 생성

2. DELETE Bucket

– Bucket 삭제

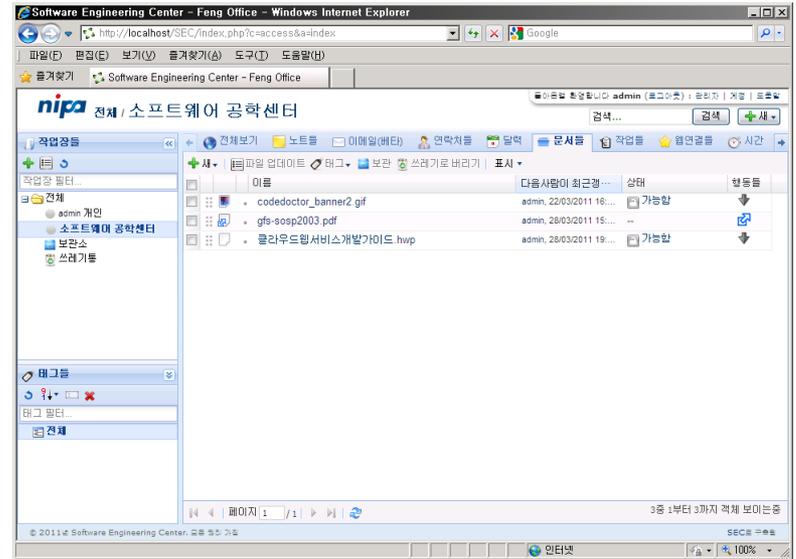


✓ 세부 HTTP 프로토콜 – S3 클래스 소스 내에 구현되어있음

1. 그룹웨어 예제



fengoffice 로그인 화면



파일 보관소 저장기능을 '웹연결' 로 연결

- ✓ 파일 저장만 AWS S3 로 변경
- ✓ 내부 소스 분석이 우선 - 가상화 레이어 필요

1. AWS Access Key

– AKIAINNZZ3BK YVITWPOA

2. AWS Secret Key

– BIBRTIfjExVqk1870nmnTEHtVMj/SpojDblsZzA

감사합니다

