

# GLORY 프로젝트: 클라우드 컴퓨팅 연구 사례

한국전자통신연구원 | 민욱기 · 김학영\* · 김영균\* · 이미영\* · 김성운\* · 남궁한

## 1. 클라우드 컴퓨팅 서비스와 기술

2006년 구글 CEO 회의에서 크리스토퍼 비시글리아가 에릭 슈미츠에게 클라우드 컴퓨팅 개념을 처음 제안한 이후, 2008년에 구글, 마이크로 소프트, SUN을 흡수한 오라클, HP, IBM 등의 최고 경영자들이 연이어 클라우드 컴퓨팅을 자사의 차세대 주력 상품으로 선언하면서 클라우드 컴퓨팅에 대한 열기는 가히 빅뱅처럼 일어났다. 우리나라에서도 2008년과 2009년의 준비과정을 거쳐 2010년부터 통신 사업자를 중심으로 클라우드 컴퓨팅 서비스가 본격화되고 있다[1-4].

서비스 기술에 치중되어 있는 우리나라 기술의 현실로 볼 때, 당분간 외산 솔루션 위주의 클라우드 컴퓨팅 서비스는 불가피한 것으로 보이지만 클라우드 컴퓨팅 역사 자체가 길지 않은 기술로 본다면 어느 정도의 기술 투자가 있다면 국내 솔루션 확보가 가능할 것으로 전망된다.

그러나 클라우드 컴퓨팅 서비스의 역사가 짧다고 기술 역사까지 짧다고 볼 수는 없다. 클라우드 컴퓨팅 서비스 기술들은 오랫동안 축적된 기술에 ‘서비스로 제공’하는 기술 부분이 더해져 완성되었기 때문이다. 예를 들어, 데이터베이스 인터페이스를 앱엔진으로 제공하는 구글의 Bigtable 인터페이스의 경우, Bigtable 자체는 이미 구글 검색 서비스를 위하여 구글의 태생부터 존재하였으며 앱엔진에서는 사용자들이 사용할 수 있는 온라인 인터페이스가 추가된 것이다[5-7]. 결국 클라우드 컴퓨팅 서비스 표면에 나타난 인터페이스가 중요한 것이 아니고, 그 인터페이스를 가능케 하는 인프라 기술이 중요함을 간과해서는 안 된다.

클라우드 컴퓨팅 서비스를 어느 서비스 형태로 할 것인지에 따라 확보해야 할 기술이 다양해진다. 가령,

응용이나 서비스를 제공하는 SaaS(Software as a Service)의 경우는 제공하는 응용 자체의 효용성과 서버 운영의 경제성을 위한 여러 사용자가 하나의 응용 인스턴스를 사용할 수 있도록 하는 멀티 테넌시 기능, 사용자의 활용 정도에 따라 서버 자원을 신축적으로 제공하는 확장성 등이 주요 기술이 된다[8,9]. 응용 개발 환경을 제공하는 PaaS(Platform as a Service)의 경우는 사용하기 편리함이나 풍부한 응용 인터페이스를 제

표 1 서비스 별 우수성 판단 기준

서비스 종류	판단 기준	설명
SaaS (Software as a Service)	멀티 테넌시	서버 상의 응용/서비스의 단일 인스턴스를 여러 사용자들이 공유할 수 있도록 함
	확장성	응용이나 서비스의 자원 사용 상태에 따라 동적으로 시스템 자원을 할당해 줌
	단말 호환성	SaaS로 제공되는 응용에 대하여 다양한 단말에서 사용 가능하도록 함
	응용/서비스 우수성	SaaS에서 제공되는 응용/서비스 자체의 효용도가 높아야 함
PaaS (Platform as a Service)	편리성	응용/서비스를 생성하는 프로그래밍 인터페이스 사용이 용이해야 함
	독립성	가급적 응용/서비스를 생성하는 플랫폼에 독립적이어야 함(현재는 대부분 종속적)
	보안성	생성되는 응용/서비스 사용에 대한 서비스 보안이 제공되어야 함
	다양성	사용자들이 API로 활용할 수 있는 함수들이 다양할수록 유리
IaaS (Infrastructure as a Service)	가용성	서버 장애 시간의 최소화
	확장성	자원을 서비스로 제공하기 위한 컴퓨팅 자원의 대규모성을 이루기 위한 기술
	안정성	서버 상에 존재하는 사용자 가상 머신, 데이터들에 대한 무손실 대책
	활용률	서버 자원의 적절한 사용으로 서버 활용률 극대화 → 서버 자원 최소 운영
	그린화	데이터 센터의 전력 절감 방안

\* 중신회원

† 본 연구는 2007년 지식경제부(구, 정보통신부)와 한국산업평가관리원에서 SW 플래그쉽 프로젝트로 추진한 “저비용 대규모 글로벌 인터넷 서비스 솔루션 개발” 과제로 2011까지 수행 예정입니다.

공하는지가 주요 척도가 된다. IaaS의 경우는 서버의 규모성, 확장성, 가용성 등이 주요 이슈가 된다. 이러한 서비스 우수성 판단 기준을 정리하면 표 1과 같다.

GLORY 프로젝트에서 확보하고 있는 기술들은 PaaS와 IaaS에 걸쳐 있다. GLORY의 대용량 데이터베이스 기술과 대규모 분산 처리 기술, 보안 기술 등이 PaaS에 활용될 수 있으며, 클러스터 관리 기술, 클라우드 컴퓨트 및 스토리지 기술, 분산 파일 시스템 기술 등이 IaaS로 활용될 수 있다. 또한 데이터센터의 전력절감을 위한 기술도 포함하고 있어 클라우드 컴퓨팅 서비스를 구축하는 데 유용할 것으로 본다.

2장에서는 GLORY 프로젝트의 개요와 GLORY 구성 기술들을 살펴보고, 3장에서는 GLORY 프로젝트의 세부 기술 내용을 그리고 4장에서는 GLORY 기술을 활용하여 국제적 컨소시엄으로 구축한 오픈 테스트베드에 대하여 소개한다.

## 2. GLORY 프로젝트

### 1.1 GLORY 개요

GLORY(GLOBAL Resource management sYstem for future internet service)는 클라우드 컴퓨팅이 탄생하기 이전인 2007년 정보통신부의 차세대 SW 산업을 위한 플래그쉽 과제로 출발하였다. 2007년 당시 국내에서도 네이버, 다음을 선두로 하는 포털 사이트들이 이미 존재하였으나, 구글, 야후 등과 같이 글로벌 경쟁력을 갖춘 기업은 없었다. GLORY 프로젝트는 글로벌 경쟁력을 가질 수 있는 국내 대규모 인터넷 서버 기술을 획득하는 것이 목표였다. GLORY 프로젝트 시작 시점에서 클라우드 컴퓨팅이라는 용어는 존재하지 않았지만, 중소기업이 하드웨어 인프라 자원을 구축하고 운영하는 것이 어렵고 초기 비용이 많이 소요되므로 이것들을 유틸리티 서비스로 제공하고자 하는 서버 솔루션은 포함하고 있었다. 이로 인하여 중소기업은 하

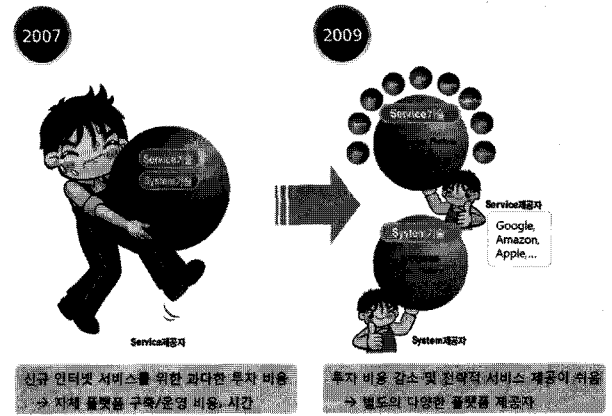


그림 1 GLORY 프로젝트에서 예견한 서비스 환경 변화

드웨어나 시스템 SW와 같은 난이도 높은 IT 인력을 배치하거나 IDC 공간을 확보할 필요 없이 대규모 서버 솔루션을 가진 업체에서 제공하는 서비스를 활용하여 상위 SW만을 탑재 운영할 수 있도록 하고자 하였다. 결국 개념적으로 클라우드 컴퓨팅이 가지는 목적과 동일하였다.

그림 1은 2007년 GLORY 프로젝트 생성 당시에 정의한 서비스 환경 변화를 예견한 그림이다. 예견처럼 2007년 말, 2008년 초에 클라우드 컴퓨팅 빅뱅이 일어났다.

GLORY 프로젝트에서 추구하는 것은 일반적인 저가 PC형 서버를 클러스터로 연결하여 저비용으로 대규모 서버 플랫폼을 구축하는 것이다. 우리나라는 웹 응용 서비스 기술로는 세계적인 경쟁력을 보유하고 있으면서도 세계 시장 진출이 부진하다. 글로벌 시장 경험도 부족하지만 저비용 분산 플랫폼, 플랫폼 확장성, 자원 유틸리티화 등과 같은 기술적인 면이 부족한 것으로 판단된다. 이에 GLORY 프로젝트에서는 저비용 분산 플랫폼을 구축하는 것을 목표로 하였다.

GLORY 프로젝트에서 저비용을 이루기 위하여 노드 당 백만원 이하의 일반적인 상용 PC급 서버를 사용하므로 인하여 구성 노드들이 빈번하게 다뤄질 수

표 2 기술적 이슈에 대한 GLORY:Google 비교

기술적 이슈		Google	GLORY
저비용		수만대 단위 주문형 Entry 수준 서버 노드	100만원 이하 PC급 노드
규모성	확장성 (Scalability)	연결노드: 1천~2천 노드/사이트 최대 저장 용량: 1015 bytes 최대 파일 수: 수천만 개(Hadoop)	연결노드: 1만 노드/사이트 최대 저장 용량: 1015 bytes 최대 파일 수: 수억 개
	가용성 (Availability)	99.999% (시스템 가용성 측면)	99.999% (시스템 가용성 측면)
	저소비 전력	12V DC 입력 (10~15% 절감)	12V or 48V DC 입력 (10~20% 절감)
활용률 (Utilization)		Unknown	45~50% (서비스기반 자원 동적 할당)

있는 것을 감안해야 한다. 따라서 구성 노드에서 발생하는 오류를 감지하고 이에 대비하는 기술들이 ‘고가용성’(High Availability) 이슈로 존재한다. 일반 저사양 PC들로 슈퍼컴퓨터 이상의 성능을 얻는 데이터 센터 효과를 얻기 위해 대규모 클러스터를 가능케 하는 ‘고확장성’(Scalability)와 ‘고활용성’(High Utilization), 그리고 전력절감을 위한 ‘저소비전력’(Low Power Consumption) 이슈들이 존재한다. 이러한 이슈들을 구글 서비스 플랫폼[10]과 비교해보면 다음 표 2와 같다.

## 1.2 GLORY 구성

GLORY 구성은 크게 동영상 기반 인터넷 서비스 구축, 동영상 서비스 구축에 필요한 글로벌 인터넷 서비스 솔루션 개발, 서비스 및 솔루션에 대한 시험 환경 구축으로 분류할 수 있으며, 글로벌 인터넷 서비스 솔루션 구성 요소들은 그림 2와 같이 다음과 같은 기술들로 구성되어 있다. 그리고 이 구성요소에 더하여 클라우드를 지원하기 위한 클라우드 인프라 기술 요소가 있다. 클라우드 인프라 구성 기술로는 클라우드 스토리지 기술과 클라우드 컴퓨트(CPU, 메모리, 운영체제) 기술이 포함된다.

### 1) 동영상 기반 인터넷 서비스 솔루션

- 동영상 서비스 솔루션: 동영상 분배, 전송 등을 지원하는 동영상 서비스 솔루션으로 네트워크 대역폭을 고려한 동영상 서비스 솔루션을 클러스터 기반 서비스 플랫폼과 연동하여 개발
- 동영상 관리 컴포넌트: 동영상 기반 인터넷 서비스를 위해 콘텐츠 및 콘텐츠 메타 데이터를 생성, 저장, 검색 등을 지원하는 솔루션으로 클러스터

기본 서비스 플랫폼 위에서 개발

- 대규모 동영상 메타데이터 저장 관리 S/W 개발
- 내용 기반 동영상 검색 S/W 개발
- 인터넷 서비스 공통 컴포넌트: 과금, 웹 서비스 등에 필요한 공통 미들웨어(웹 서버, 애플리케이션 서버 등)는 포털 업체에서 보유하고 있는 기존 솔루션 혹은 오픈 소스를 클러스터 기반 서비스 플랫폼에 탑재

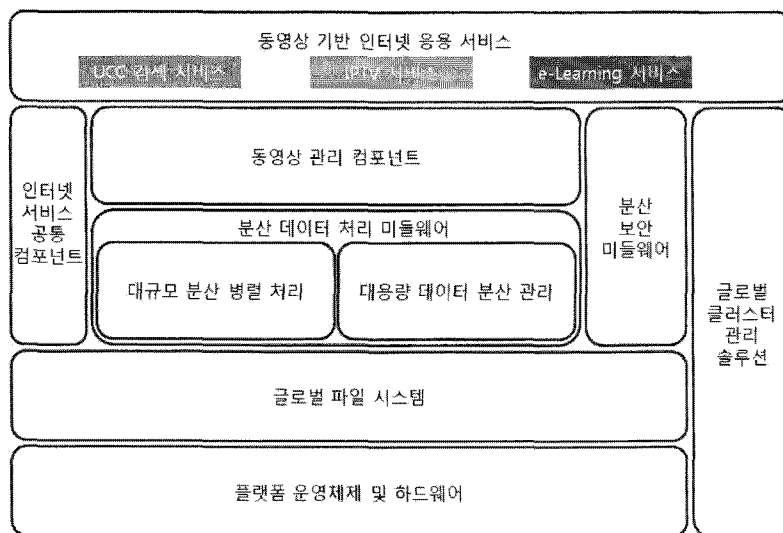
### 2) 분산 데이터 처리 미들웨어

대규모의 노드들로 구성된 클러스터 시스템 환경에서 대량의 데이터를 이용한 인터넷 서비스 개발을 효율적으로 할 수 있도록 분산 저장 관리하고, 이들 대량의 데이터에 대한 분산 병렬 처리를 수행함으로써 고성능 인터넷 서비스를 지원하기 위한 미들웨어 개발

- 대용량 데이터 분산 관리 S/W
- 인터넷 서비스를 위한 텍스트 데이터, 멀티미디어 데이터 저장 관리 기능
- 데이터 분산 배치 관리 기능
- 데이터 접근 및 복구 관리 기능
- 대규모 분산 병렬 처리 S/W
- MapReduce 모델 기반의 병렬 처리를 위한 작업 분해 및 병합 기능
- 분산 작업 스케줄링 정책 제공 및 데이터 위치 기반 부하 분산 기능
- 분산 작업 가용성을 위한 노드 고장에 대한 분산 처리 프로세스 자동 복구

### 3) 분산 보안 미들웨어

장치/커널 인증 및 사용자/서비스 인증을 통한 안전



[글로벌 인터넷 서비스 솔루션 구성도]  
그림 2 GLORY 시스템 구성 기술들

한 글로벌 인터넷 서비스 플랫폼 환경을 제공하는 분산 보안 미들웨어 기술 개발

- 신뢰 플랫폼 제공을 위한 장치/커널 인증 기능
- 트러스티드 부팅을 통한 커널 이미지 인증
- S/W 서명 및 검증을 통한 소프트웨어 다운로드 인증
- 노드간 상호 인증
- 다중 신뢰 도메인 간 안전한 ID 공유를 위한 사용자/서비스 인증 기능
- ID 연동(Federation)
- 익명/가명 인증

#### 4) 글로벌 파일 시스템

글로벌 인터넷 서비스에서 사용되는 대용량의 데이터 저장 및 처리 요구를 만족시키기 위하여 서비스 부하(work load)에 따른 데이터 입출력을 처리하고, 고수준의 데이터 가용성 및 신뢰성을 보장하며, 고 확장성 클러스터에서 데이터 공유를 지원하는 분산 파일 시스템 개발

- 파일 시스템 메타데이터 분산 관리 기능
- 메타데이터 분산 관리 기능(메타데이터 배치, 메타데이터 캐싱, 메타데이터 핫스팟 회피 흐름 제어)
- 메타데이터 서버 고장 고속 회복 및 노드 고장 탐지 및 복구
- 파일 데이터 분산 저장 및 복제 관리 기능
- 데이터 저장 서버 볼륨 구성 기능
- 데이터 가용성 지원 기능
- 데이터 고속 입출력 기능

#### 5) 글로벌 클러스터 관리 솔루션

저가 노드를 사용하는 100만 대급의 대규모 다중 데이터 센터를 운용 관리하기 위한 계층적 클러스터 시스템의 구성 및 자동 관리 기능을 제공하여 온라인 확장성 및 가용성을 제공하는 솔루션 개발

- 대규모 클러스터의 서비스 분산(orchestration)을 위한 클러스터 구성, 노드 상태 관리 및 분산 부하 관리
- 다중 데이터 센터를 지원하는 계층적 클러스터 구성
- 클러스터 구성 정보 및 정책 관리
- 클러스터 노드들의 가용 상태(aliveness) 관리
- 노드 자동 재구성
- 서비스 노드들의 부하를 고려한 상황 적응형 서비스 관리
- 대규모 자동 프로비저닝
- 시스템 자동 인식 기능

· S/W 스트리밍 기반 OS 및 소프트웨어 자동 설치 및 자동 업그레이드

- 자동 구성 기능
- 대규모 분산 자원 단일 포인트 관리를 통한 모니터링
- 웹 기반 분산 자원 통합 관리 도구
- 클러스터 시스템 내의 모든 노드들의 자원 감시

#### 6) 플랫폼 운영체제 및 하드웨어

저가격/저전력/대규모 인터넷 서비스 클러스터 시스템을 위한 상용 하드웨어 플랫폼 기반의 기능 집약형 노드 개발과 이를 이용한 대규모 클러스터 시스템 구축 기술 개발

- 리눅스 기반 저전력 H/W 플랫폼
- PC급 저비용 노드 기반의 저전력 소모형 웹박스
- 노드 자원 및 전원 관리
- 무병목 네트워크 트래픽 처리
- 공개 SW 기반의 운영체제
- 클러스터 컴퓨팅 플랫폼 구축을 위한 네트워크 위상
- 클러스터링을 위한 네트워크 위상
- 단일 데이터 센터 내 클러스터 구성

#### 7) 저비용 클라우드 스토리지 SW

클라우드 스토리지 서비스를 위한 여러 사용자/서비스별 대용량 파일 저장 관리 및 파일 입출력 요구를 만족시키는 다중테넌트(multi-tenant) 지원 스토리지 플랫폼 개발

- 다중 개인 사용자/서비스 지원 다중 파일 셋(multiple file set) 기능
- 썬 프로비저닝 지원을 위한 스토리지 풀 관리 및 최적 파티셔닝
- 파일셋별 고가용성 지원
- 클라우드 데이터 보호 기능
- 서비스 및 사용자별 파일 접근 제어
- 파일 암호화
- 클라우드 스토리지 절감 기능
- M+N(RAID 6)
- 데이터 중복 제거(de-duplication)
- 파일 메타데이터 분산 관리 기능
- 멀티 메타데이터 서버 구성 및 메타데이터 분산
- 대량 파일 메타데이터 색인 및 캐싱

#### 8) 클라우드 컴퓨트 제공 SW

사용자 요구 기반 컴퓨트(CPU+Memory)를 생성, 제 공함에 있어, VM 이미지 관리 효율화, 유연성 확보, 사용자 대기 시간 최소화로 대규모 컴퓨팅 자원을 관

리, 운영하는 클라우드 컴퓨트 SW 개발

- 공개 가상화 도구 기반 이기종 VM 프로비저닝
- 리눅스/윈도우/안드로이드 등 사용자 요구 기반 VM 프로비저닝
- VM 이미지 클로닝 방식의 고속 deployment
- 능동적인 VM 스케줄러
- 사용자 별 VM 모니터링 및 관리
- VM & 노드별 상황에 따른 VM 배치, 마이그레이션
- VM의 능동적인 확장
- Deduplication 방식의 VM 이미지 관리
- 공통적인 커널 이미지와 사용자 이미지 분리 저장
- VM 이미지 풀 관리

### 3. GLORY 세부 기술

#### 3.1 대규모 클러스터 관리 SW

글로벌 규모를 지원하기 위한 클러스터 관리 SW는 그림 3과 같이 클러스터 오케스트레이션, 자동 프로비저닝, 그리고 동적 서비스 관리로 구성된다.

단일 데이터 센터에서 1만대를 하나의 클러스터를 구성하기 위해서는 기존의 플랫(flat)한 구조를 사용할 수 없다. 마스터 하나를 두고 1만대를 운영하려면 마스터에 집중되는 부하를 감당하기 어렵기 때문이다. GLORY 클러스터 관리 SW에서는 그림 4와 같이 계층적으로 클러스터를 구성하였다.

계층적인 클러스터에서는 256대의 상태를 모니터링하고 오케스트레이션하는 그룹 마스터들이 존재하며 클러스터 마스터는 그룹 마스터를 모니터링하며 40개

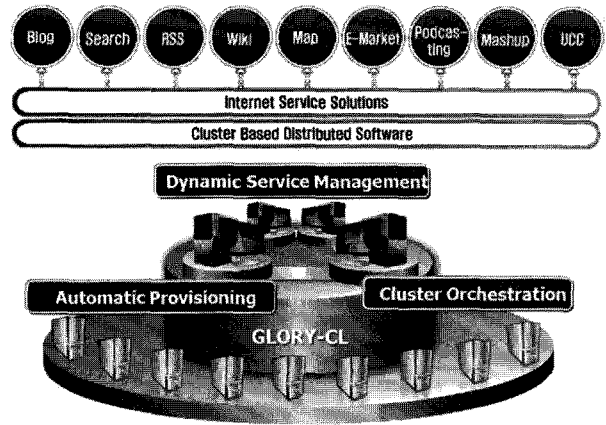


그림 3 대규모 클러스터 관리 SW 구성

의 그룹 마스터까지 둘 수 있다. 이러한 계층적인 클러스터 구성은 노드가 확장될 때 자동으로 구성될 수 있으며, 특정 노드의 장애가 발생하면 노드 성격에 따라 적절한 조치를 취한다. 가령 그룹 마스터에서 장애가 발생하면 클러스터 마스터는 동일 그룹 내의 일반 노드 중에 하나를 그룹마스터로 임명하고 이전 그룹 마스터를 fail로 셋팅한다. 클러스터 마스터에서 장애가 발생할 경우를 대비하여 클러스터 마스터는 항상 active-active 상태의 HA(High Availability) 기능을 사용하고, 클러스터에 있는 노드들을 할당해 주기 위한 풀(Pool)을 관리한다. 파일시스템이나 데이터베이스, 분산 처리, 또는 특정 서비스를 위하여 풀에서 원하는 노드 수 만큼의 자원을 요청할 수 있고, 필요에 따라 반납 처리되기도 한다. 자원 모니터링은 메시지 전송을 최소화하여 시스템 부하를 최대한 줄이도록 설계되었다. 하나의 그룹 마스터에서 256대의 노드로 확장될 경우 CPU 부하는 0.1% 미만이며 일반 노드들은 메

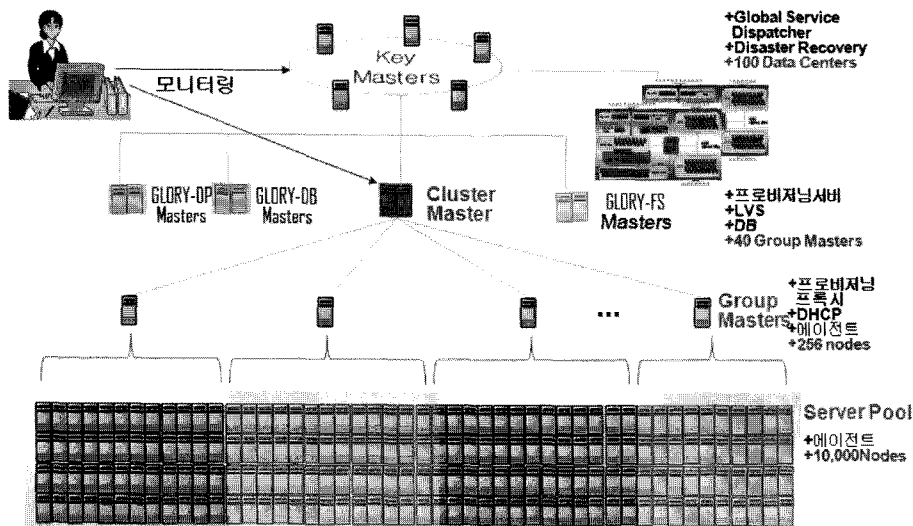


그림 4 계층적인 클러스터 구조

모리 1.6 M 정도를 그룹마스터는 2.1 M 정도를 사용한다. 그리고 네트워크 대역폭 사용도 초당 5~8 KB에 지나지 않았다.

서비스 관리 기능은 서비스 부하에 따라 서비스 노드를 할당해 주는 방식으로 클러스터의 활용률을 높여 준다. 서비스 로드 밸런서로는 LVS(Linux Virtual Server)와 L4 스위치를 모두 사용하며 특정 서비스를 등록, 제거하고 이때 서비스를 운영하는 노드수는 클러스터에서 자동적으로 할당한다. 이때 할당 단위는 하드웨어적인 머신(Physical Machine)이 되지만 이 기능을 가상화 기능과 접목하여 가상 머신(Virtual Machine) 기능으로 대입하면 동적인 확장성을 가지는 클라우드 시스템이 된다. 그러나 시장에 따라서는 가상 머신보다 하드웨어적인 머신 단위의 노드를 선호하는 경우가 있으므로 2011년 GLORY 시스템에서는 신축적인 하이브리드 머신 제공 기능으로 확대될 예정이다.

자동 프로비저닝은 초기 머신에 운영체제 및 필요한 SW를 자동으로 설치하고 설정하는 SW이다. GLORY 프로비저닝은 P2P 형태의 Bit Torrent 방식을 사용하여 256대의 시스템을 17분 만에 설치할 수 있다.

### 3.2 CCS: 클라우드 컴퓨트 SW

GLORY-CCS(Cloud Compute SW)는 가상 머신을 일반 사용자에서 서비스를 제공하기 위한 솔루션으로 클러스터 관리 SW를 기반으로 구성된다. GLORY-CL 부분은 기존의 클러스터 관리 SW이며 GLORY-CCS는 클라우드 컴퓨트를 제공하기 위한 모듈이다. GLORY-CCS는 공개 소프트웨어로 제공되는 가상화 도구(KVM, Xen)[11,12]들과 Tashi[13], libvirt와 같은 가상화 관리 도구를 기반으로 퍼블릭 클라우드를 제공하기 위하여 필요한 기술들을 추가 지원한다.

GLORY-CCS의 가상 머신 프로비저닝은 사용자 요구 사항을 기반으로 코어 수, 메모리 크기, 운영체제 선택에 따라 가상 머신을 온라인상에서 제공해 주고, 사용자는 할당받은 가상 머신을 모니터링하거나 사용할 수 있다.

가상 머신을 제공할 때, 현재 시스템 상황에 따라 적절한 노드를 선택하는 가상 머신 스케줄러 기능을 제공한다. 서비스 운영자들은 서비스가 실행되고 있는 가상 머신의 마이그레이션을 최소화하는 것과 서버 활용률을 극대화하는 것을 원한다. 서비스의 안정성을 고려하면 스트라이프(stripe) 방식의 순차적인 스케줄러 방식이 효율적이고, 서버의 활용률을 최대화하려면 팩킹(packing) 방식으로 운영하는 것이 유리하다. 그러나 시스템 활용률과 서비스 안정성을 동시에 고려하여 Min-Max 알고리즘으로 가상 머신 할당 노드를 결정한다. Min-Max 알고리즘은 새롭게 생성되는 가상 머신의 리소스가 현재 노드의 리소스 상황과 합쳐서 Min-Max 범위 안에 속하는 노드 중 가상 높은 노드를 선택한다. 이로 인하여 서버의 활용률을 Max값에 근접하는 활용률을 얻고, Min 값 이하 노드들의 전원은 차단하여 전력절감의 효과를 얻을 수 있다.

사용자들이 많아지므로 인하여 1~16 M에 이르는 사용자 가상 머신의 관리 또한 해결해야 할 문제가 된다. 일반적인 4 M 가상머신을 가정하면 사용자가 백만 사용자에게 이르면 사용자 데이터를 제외한 가상머신 공간만으로도 페타 규모의 디스크 용량을 요구하게 된다. 이를 감축하기 위해서 GLORY-CCS에서는 De-duplication 방법을 사용하여 가상 머신 공간을 감축한다. 커널을 기준으로 그룹화하면 30% 이상의 저장 공간을 감축할 수 있다.

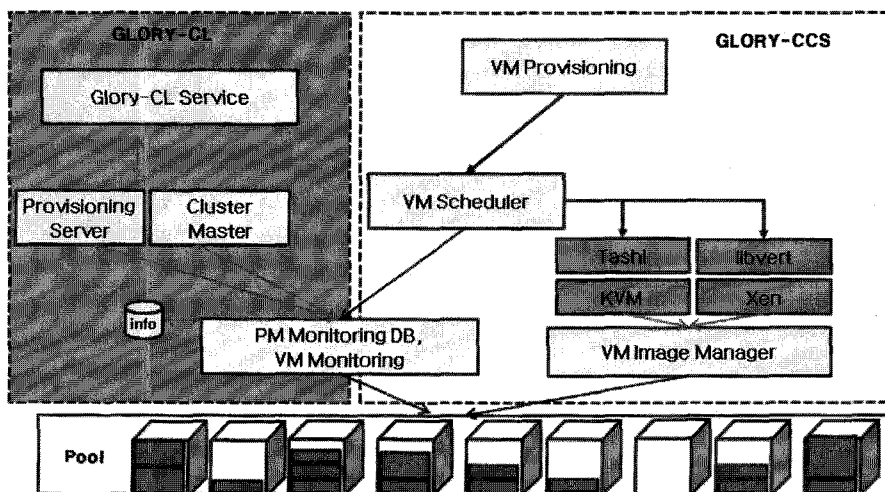


그림 5 GLORY-CCS와 클러스터 상관관계

### 3.3 글로리 파일 시스템

글로리 파일 시스템(GLORY-FS: GLORY File System)은 분산된 PC급의 저규모 서버들의 디스크를 단일 파일 시스템으로 운영하여 TCO를 최소화시킨 분산 파일 시스템이다.

10테라바이트(500GB disk × 20개)의 저장 공간을 가진 노드들을 200개까지 하나로 묶어 1.8 페타바이트 저장 공간을 제공할 수 있으며, 200개 노드 범위 내에서 150 GB/초의 입출력 성능을 얻을 수 있다. 또한, 글로리 파일 시스템에 저장된 데이터들은 복수 개로 카피되어 저장되며 노드 오류를 관리함으로써 고가용성을 보장한다.

MDS(Meta Data Server)의 초당 동시 처리 연산 성능을 최적화하고 하이브리드 캐시를 구현하여 캐시를 사용하지 않았을 때 보다 1.5배의 동시처리 연산 성능을 얻었으며, 응용 서비스 규모에 최적화된 가상 메타데이터 스토리지도 구현되어 있다. GLORY 파일 시스템의 클라이언트로는 리눅스는 물론이고 윈도우즈 계열까지 가능하다.

### 3.4 CSS: 클라우드 스토리지 SW

GLORY-CSS(Cloud Storage SW)는 페타 바이트 저장 공간을 지원하는 GLORY 파일 시스템을 기반으로 클라우드 스토리지 서비스를 제공하기 위한 SW 스택이다.

그림 7과 같이 GLORY-CSS에서는 사용자들에게 개별적인 볼륨을 제공하고 볼륨별 쿼터값을 설정하는 방식으로 다중 테넌트를 지원할 수 있다. 개인별 스토리지 쿼터는 한도값 설정을 변경할 수 있고 현재 사용량을 조회할 수 있는 사용자 서비스 기능과 신속성을 제공한다. 가상 머신에서 연동되는 클라우드 스토리지 서비스를 위하여 가상 머신 파일 입출력 특성을 최적화시킨 데이터 통신 프로토콜을 구현하였으며, VMWare와 시트릭스와 같은 상용 가상화 제품과의 연동을 지원하기 위하여 NFS 프로토콜도 개발하였다. 또한 데이터 서버 장애 발생 시 복제 트래픽 증가로 인한 사용자 서비스 장애를 유발을 회피하기 위하여 서비스 네트워크와 복제 네트워크를 분리하는 방식으로 서비스 장애를 최소화 하였다.

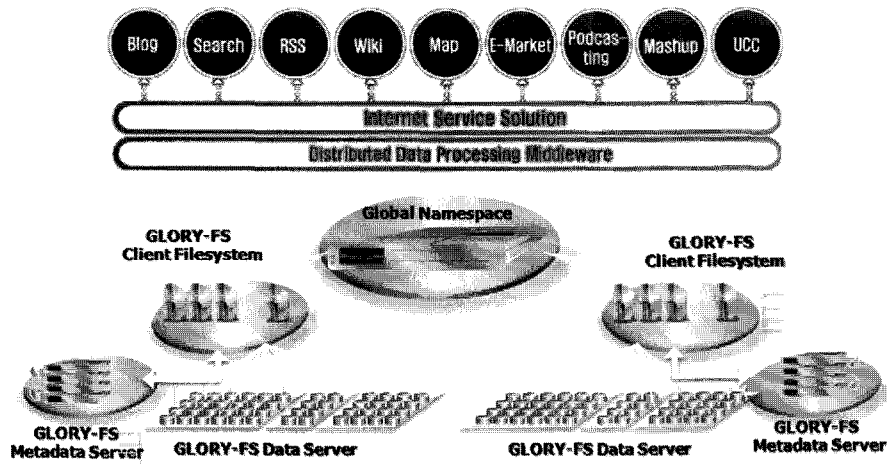


그림 6 GLORY-FS 구성 요소

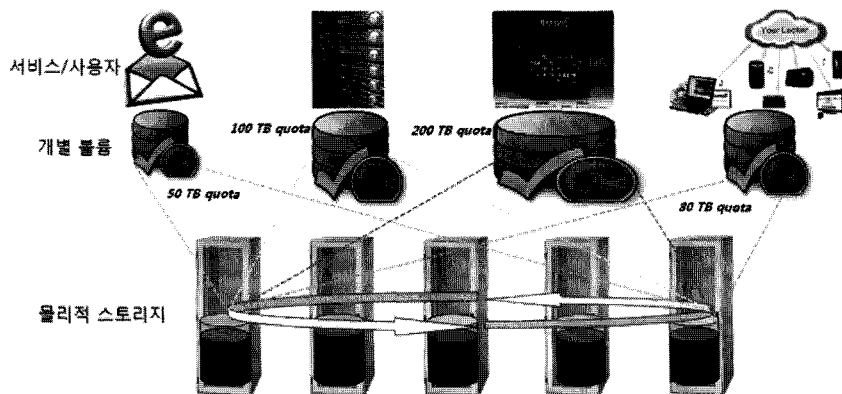


그림 7 GLORY-CSS의 개인 볼륨 제공

### 3.5 대용량 데이터 분산 관리 시스템

대용량 데이터 분산 관리 시스템(GLORY-DB : GLORY DataBase)는 대규모 노드들로 구성된 클러스터 파일 시스템을 기반으로 대량의 데이터를 분산 저장 관리하는 데이터베이스 시스템이다.

GLORY-DB는 Hadoop의 Hbase[15]를 확장한 시스템으로 key-value 기반의 저장 관리, 컬럼 그룹 기반으로 데이터 저장, 키 범위 기반 데이터 분산 관리 등의 기능으로 인터넷 서비스에서 요구하는 고확장성의 데이터 관리 기능을 제공한다. 또한 SQL과 유사한 질의어를 지원하므로 응용 개발의 편리성을 제공하고 행 스캐너와 셀 스캐너 검색 기능이 모두 가능하다.

GLORY-DB는 데이터 관리 서비스의 고가용성을 위하여 노드, 네트워크 오류에 대한 GLORY-DB 복구 기능을 제공한다. 주기적인 스냅샷 저장 관리를 이용하

여 마스터 서버의 오류 복구를 제공하며, 각 파티션 서버의 오류에 대처하기 위해서는 로깅을 수행, 이를 이용하여 장애를 복구한다.

GLORY-DB는 텍스트 데이터 관리 외에도 동영상 내용기반 검색 서비스가 가능하도록 2단계 트리와 시그니처 기반으로 분산 고차원 인덱스 관리 및 kNN 유사 검색 기능을 제공한다. 분산 고차원 인덱스 관리 기능은 특징 벡터 데이터 량에 대한 확장성을 제공하면서 평균 95%의 검색 정확도를 제공한다.

GLORY-DB는 웹 포털 서비스, UCC 동영상 서비스 구축 등에 필요한 대용량 데이터 관리에 유용하다.

### 3.6 대규모 데이터 분산 병렬 처리 시스템

대규모 데이터 분산 병렬 처리 시스템(GLORY-DP : Glory Distributed Processing)은 수천대 노드 규모에서

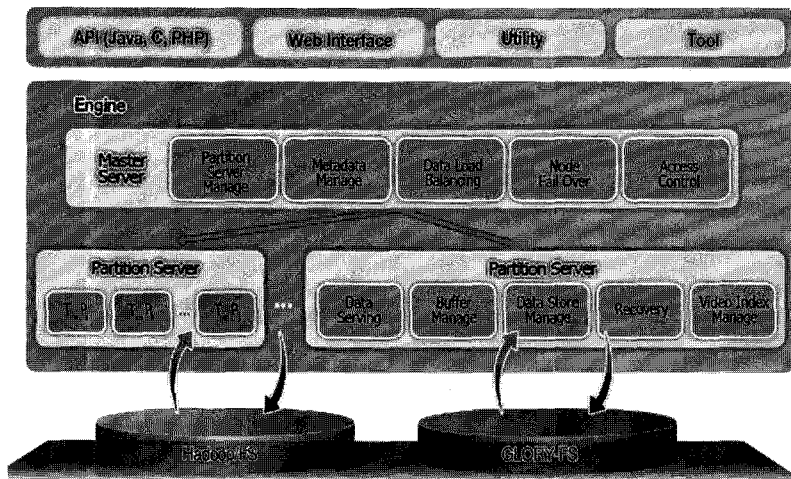


그림 8 GLORY-DB의 구조

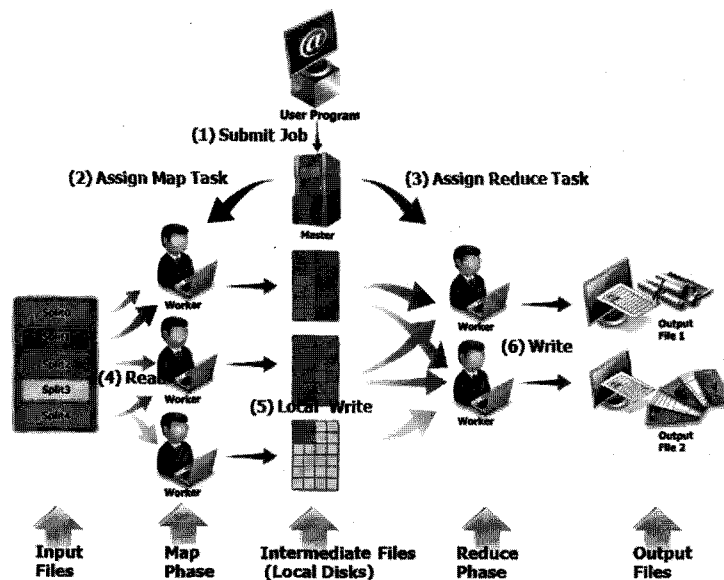


그림 9 GLORY-DP에서 작업처리 과정



대규모 데이터 처리를 병렬로 수행할 수 있는 대규모 분산 병렬 처리 시스템이다.

GLORY-DP는 Hadoop의 MapReduce[16]를 확장한 시스템으로, 맵리듀스 (Map&Reduce)[17] 프로그래밍 모델을 기반으로 분산 병렬 처리 응용을 쉽게 개발할 수 있는 환경을 제공한다. 맵리듀스 프로그래밍 모델은 그림 9와 같이 대규모 데이터를 분할하여 동시에 Map 업무를 처리하고 수행 결과를 수집하여 Reduce 업무를 처리하여 최종 결과 파일을 생성하는 방식으로 이루어진다.

GLORY-DP는 입력 데이터 소스 시스템으로 GLORY-FS와 HDFS(Hadoop File System)[18], 그리고 GLORY-DB가 사용 가능하다.

GLORY-DP는 사용 환경에 맞게 스케줄링 정책을 선택할 수 있도록 DP(Direct Priority), EDF(Earliest Deadline First), LJF(Longest Job First), SJF(Shortest Job First), FIFO(First In First Output), RR(Round Robin), WRR(Weighted Round Robin), FQ(Fair Queuing), FS(Fair Share) 등과 같이 다양한 스케줄링 정책을 제공하며 혼합 사용이 가능하다. 또한 스케줄 정책을 온라인으로 변경 가능하도록 하여 유연성을 향상시켰다.

GLORY-DP는 Java API로 작성된 응용 외에도 C, C++, Perl, Python 등 다양한 프로그래밍 언어로 작성된 기존 응용 연동도 지원하며, 데이터 위치와 노드의 자원 상태를 인지한 부하 분산 기능이 제공된다. 또한 노드 장애나 네트워크 오류로 인한 GLORY-DP(Job Tracker와 Task Tracker로 구성) 장애 복구 기능을 제공하고, 타스크 수행 오류를 인지하여 자동으로 재실행 시키므로 인하여 분산 처리 작업의 고가용성을 제공한다.

GLORY-DP는 웹 페이지를 인덱싱하거나 분자 구조,

DNA 염기서열 분석 등과 같이 대규모 데이터를 처리하는 데 유용하여 데이터 중심의 클라우드 컴퓨팅 시스템에 효과적으로 사용될 수 있다.

### 3.7 저전력 하드웨어

GLORY-WB(Web Box)는 저비용, 저-TCO, 저전력을 제공하는 시스템으로 하드웨어 전력 소모량을 20% 이상을 절감하였다. 랙준위 DC 전력전달과 AC 서버를 DC 서버로 변경하는 PDB(Power Distribution Board)를 개발로 서버 전력 소모량을 절감하였다. 따라서 서버 소모전력이 낮아지고 이로 인하여 데이터 센터 냉각 비용도 절감하여 데이터 센터 전체 전력 소비량을 낮출 수 있다.

GLORY RPSU(Rack Power Supply Unit)는 서버랙에 87%의 고효율 DC 전력을 공급하여 높은 전력 효율성을 가지고 있다. RPSU에서는 실시간 전력 및 온도 습도를 네트워크를 통해 모니터링 할 수 있으며 이를 관리하는 RPMS(Rack Power Management SW)를 제공한다. 그림 10은 RPSU를 통한 랙전력과 환경을 모니터링하고 컨트롤하는 RPMS 구성을 보여준다.

## 4. 오픈 시러스 테스트베드

오픈 시러스[19]는 미국의 HP, 인텔, 야후 주관으로 클라우드 컴퓨팅을 위한 오픈 테스트베드 컨소시엄으로 현재는 HP, 인텔, 야후 이외에도 독일의 KIT, 중국의 차이나 모바일과 차이나 텔레콤, 러시아 한림원, 싱가포르의 DIA, 말레이시아의 MIMOS, 미국의 일리노이 대학, CMU 등과 함께 ETRI도 참여하고 있다.

오픈 시러스는 각 기관의 테스트베드를 해당국의 사용자들을 대상으로 클라우드 컴퓨팅을 사용할 수 있도록 하며, 여기서 얻어지는 경험을 1년에 두차례 열

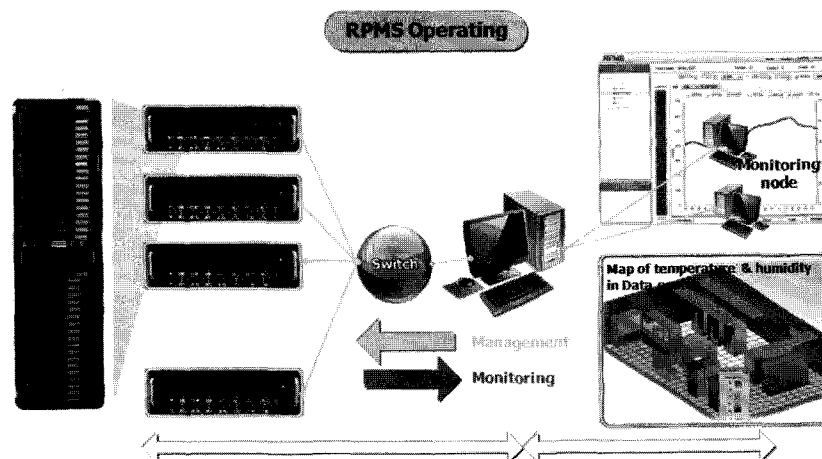


그림 10 저전력 하드웨어 시스템 구성

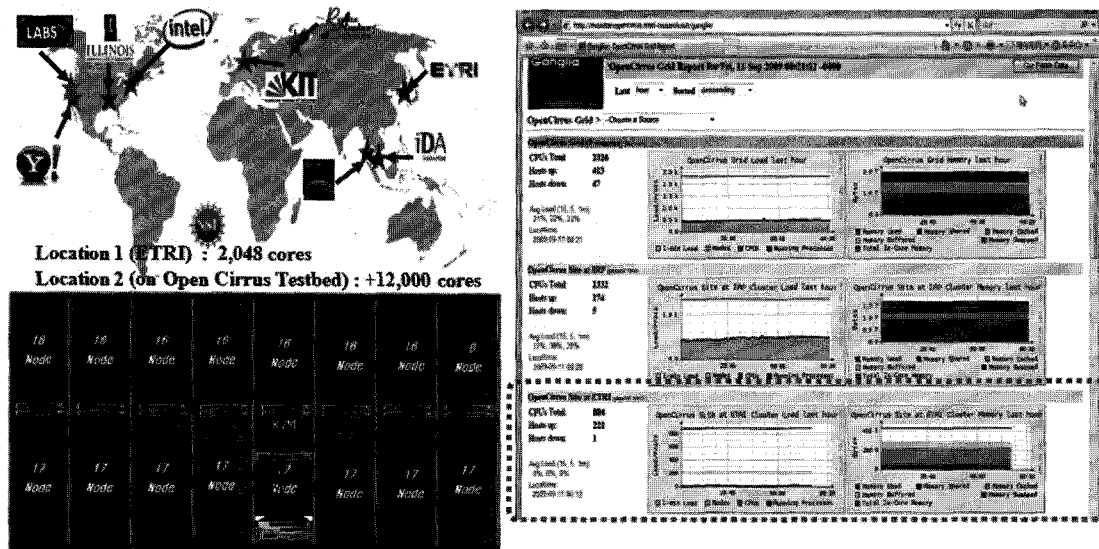


그림 11 오픈 시러스 ETRI 테스트베드

리는 오픈 시러스 씨밋에서 공유한다.

에트리에서는 오픈 시러스 테스트베드를 위하여 현재 512대 2000 코어를 제공하고 있으며 각 시스템은 4-core 프로세서와 2GB 메인 메모리를 가지고 있다. 에트리 오픈 시러스 테스트 베드는 공개 소프트웨어들로 구성된 오픈시러스 공통 플랫폼과 GLORY 프로젝트 결과물을 동시에 설치하였으며, 사용자의 요구에 따라 플랫폼을 사용할 수 있는 기회를 부여한다. 오픈 시러스 공통 플랫폼으로 사용하는 공개 소프트웨어는 HDFS, Hadoop MapReduce, Xen, KVM, Tashi, Zoni, Ganglia 등이며 Ganglia 모니터링 툴을 사용하여 그림 11과 같이 각 가입 사이트들의 현재 상태를 모니터링할 수 있다.

## 5. 결론

GLORY 프로젝트는 2007년 시작으로 1단계(2007~2009년) 개발로 저비용 대규모 분산 컴퓨팅 기술을 개발 완료하고, 현재 2단계(2010~2011년)에서 분산 컴퓨팅 기술을 활용하여 인프라 자원을 서비스로 제공하기 위한 클라우드 컴퓨트 SW와 클라우드 스토리지 SW 기술 개발을 진행 중이다.

GLORY 프로젝트의 1단계 결과물 중 GLORY-FS은 현재 케이타하이텔(주)의 파란 인터넷 서비스와 클루넷 CSS(Cloud Storage Service)에 활용되었으며, GLORY-WB 기술은 (주)대영정보통신 등에 기술이전 되어 IDC 및 공용 PC 시설로 사업화하고 있다.

2단계 결과물인 클라우드 컴퓨트 및 스토리지 서비스 제공 SW가 개발 완료되면 국내 클라우드 인프라

기술을 한 단계 끌어올리는 견인차 역할을 하게 될 것이다.

## 참고문헌

- [1] 민옥기, 이미영, 허성진, 김창수, 흰히 보이는 클라우드 컴퓨팅, 전자신문사, 2009.
- [2] 정제호, “클라우드 컴퓨팅의 현재와 미래, 그리고 시장 전략”, 2008년 10월
- [3] 신현석, “마이크로소프트와 클라우드 컴퓨팅 MS 클라우드 전략의 ‘코어’, 윈도우 애저”, 마이크로소프트 웨어, 2009년 1월, pp 160~163
- [4] Amazon, “Amazon Web Services: Overview of Security Process”, <http://aws.amazon.com white paper>, September 2008
- [5] FayChang, et.el, “Bigtable: A Distributed Storage System for Structured Data”, 7th USENIX Symposium on OSDI, 2006
- [6] George Lawton, “Developing Software Online with Platform-as-a-Service Technology”, Computer, June 2008
- [7] Ray Valdes, “Google App Engine Goes Up Against Amazon Web Services”, Gartner, April 2008
- [8] 세일즈포스닷컴, “Salesforce 마케팅”, <http://salesforce.com>
- [9] KIPA, “SaaS 대표주자, Salesforce.com의 성장세 분석”, 2007년 11월
- [10] SanjayGhemawat, et.el. “The Google File System”, 19th ACM Symposium on Operating Systems Principles, 2003

- [11] Avi Kivity et.al, Kvm:the Linux Virtual Machine Monitor, Linux Symposium, 2007
- [12] Jin Heo, et.al, Memory Overbooking and Dynamic Control of Xen Virtual Machines in Consolidated Environments, IFIP/IEEE IM , 2009
- [13] <http://incubator.apache.org/tashi>
- [14] IBM, “가상화 기술 백서”, <http://www.ibm.com>, 2006
- [15] <http://hadoop.apache.org/index.pdf>, “Welcome to Apache Hadoop!”
- [16] <http://hadoop.apache.org/mapreduce>
- [17] JeffreyDean, SanjayGhemawat, “MapReduce: Simplified Data Processing on Large Clusters”, OSDI '04, 2004
- [18] <http://hadoop.apache.org/hdfs/>
- [19] Arutyun I. Abetisyan, 외9인, “Open Cirrus: A Global Cloud Computing Testbed”, IEEE Computer Society, 2010.4.

## 약 력



### 민 옥 기

1988 충남대학교 계산통계학 학사  
 1992 충남대학교 계산통계학 석사  
 2010 충남대학교 컴퓨터공학과 박사  
 1988~현재 한국전자통신연구원 책임연구원  
 관심분야 : 클라우드컴퓨팅, 모바일 클라우드, 자  
 원 가상화, 클러스터 관리, 미디어 전송  
 E-mail : ogmin@etri.re.kr



### 김 학 영

1983 경북대학교 전자계산 학사  
 1985 경북대학교 전자계산 석사  
 2003 충남대학교 컴퓨터공학과 박사  
 1988~현재 한국전자통신연구원 책임연구원  
 관심분야 : 클라우드컴퓨팅, 가상화, 클러스터 관  
 리, 미디어 전송  
 E-mail : h0kim@etri.re.kr



### 김 영 군

1991 전남대학교 전산통계학 학사  
 1993 전남대학교 전산통계학 석사  
 1995 전남대학교 전산통계학 박사  
 1995~현재 한국전자통신연구원 책임연구원, 저  
 장시스템연구팀 팀장  
 관심분야 : 파일 시스템, 데이터베이스, 분산 컴퓨팅 등  
 E-mail : kimyoung@etri.re.kr



### 이 미 영

1981 서울대학교 식품영양학과 학사  
 1983 서울대학교 계산통계학과 석사  
 2005 충남대학교 컴퓨터공학과 박사  
 1988~현재 한국전자통신연구원, 데이터베이스연  
 구팀 팀장  
 관심분야 : 데이터베이스 관리, 데이터 스트림 처  
 리, 분산 컴퓨팅, 클라우드 컴퓨팅  
 E-mail : mylee@etri.re.kr



### 김 성 운

2006 충남대학교 전자공학과 박사  
 1989~현재 한국전자통신연구원 책임연구원, 서  
 비플랫폼연구팀 팀장  
 관심분야 : 컴퓨터 구조, 클라우드 컴퓨팅, 저전력  
 하드웨어 플랫폼  
 E-mail : ksw@etri.re.kr



### 남 공 한

1976 고려대학교 물리학 학사  
 1982 Univ. of Columbia 컴퓨터공학 석사  
 2006 정보통신대학원대학교 박사  
 1982~1986 LG 전자 선임연구원  
 1987~현재 한국전자통신연구원, 클라우드컴퓨팅  
 연구부 부장  
 관심분야 : 클라우드 컴퓨팅, 슈퍼컴퓨팅, 분산컴퓨팅, Ad-hoc 네트워크  
 E-mail : kimyoung@etri.re.kr