



PVFS2 활용을 통한 클라우드 기반 대용량 데이터 처리 플랫폼 개발

2011. 09. 30





AGENDA

- | | |
|----|------------|
| 01 | 사업 개요 |
| 02 | 개발 목표 및 내용 |
| 03 | 적용기술 소개 |
| 04 | 개발 내용 |
| 05 | 적용 방법론 |
| 06 | 서비스 현황 |

01. 사업 개요

사업 개요

과제명

PVFS 활용을 통한 클라우드 기반 대용량 데이터 처리 플랫폼 개발

주관기관

(주) 이노그리드

총괄책임자

손호준 연구소장

참여기관/책임자

(사)한국클라우드서비스협회 / 민영기 사무국장

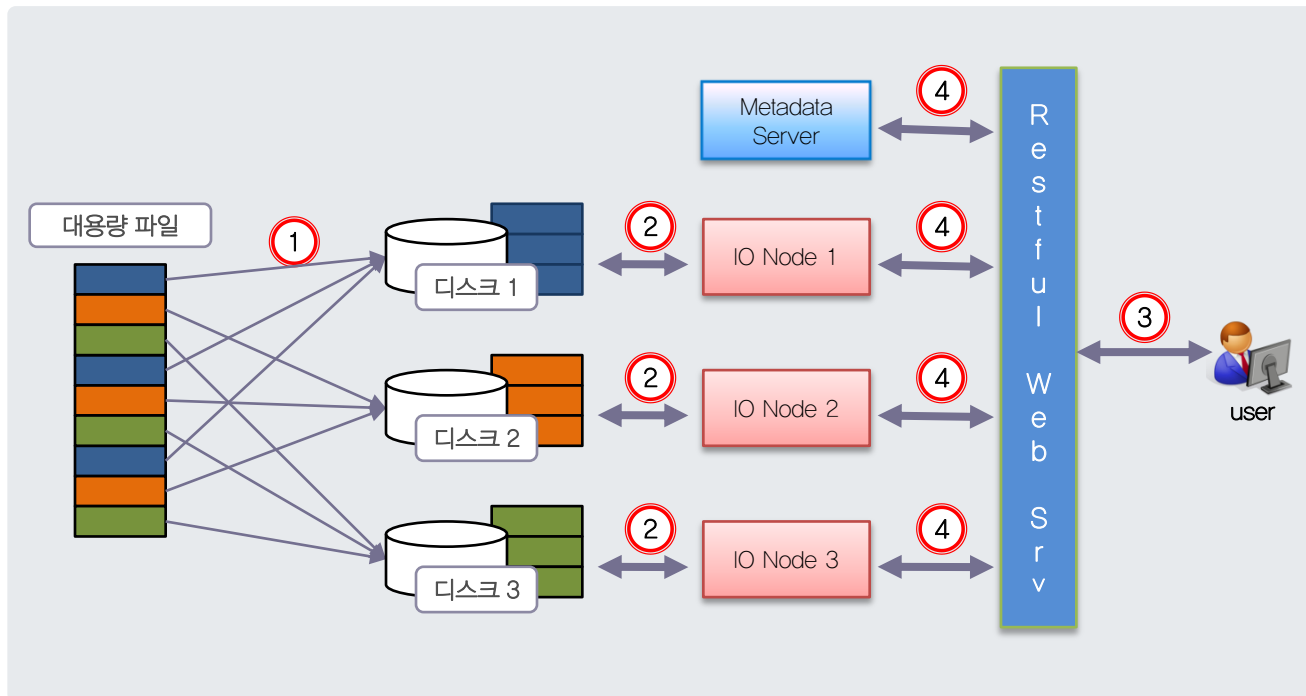
과제 수행기간

2011. 4. 25 ~ 2011. 11. 24 (7개월)

02. 개발 목표 및 내용

개발 목표 및 내용

PVFS2 활용을 통한 대용량 데이터 처리 플랫폼 개발 및 상품화



① 파일 분산 저장

② 데이터 저장 및 탐색 제어

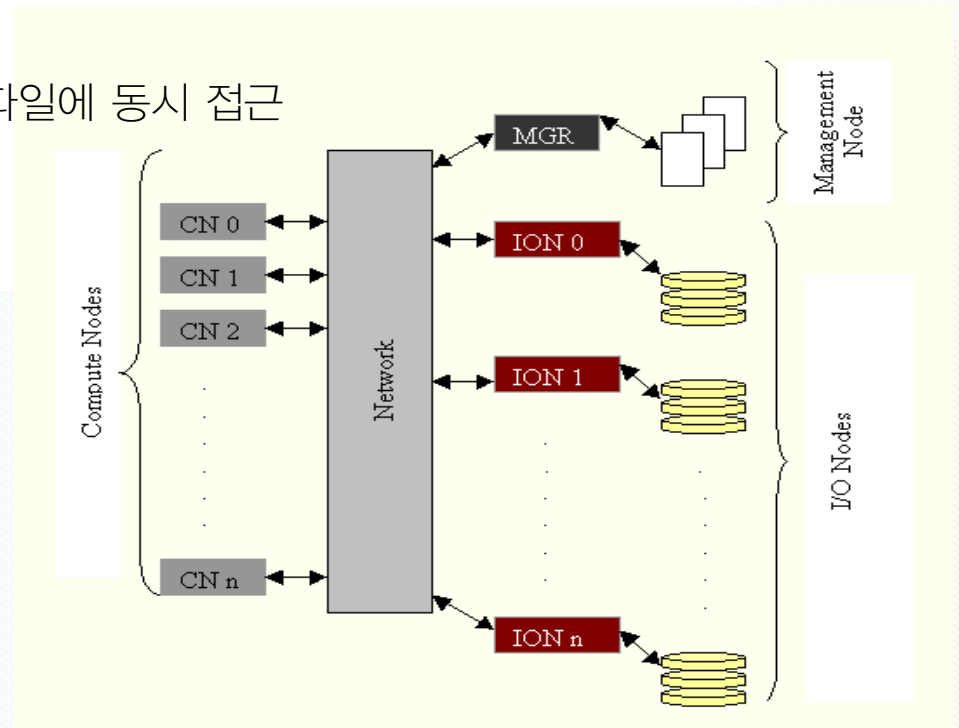
③ Restful Web Service를 통한 데이터 접근

④ 사용자 요청에 따른 분산 파일 시스템 제어

03. 적용 기술 소개

PVFS2 소개

- Parallel Virtual File System 2
- 분산 파일 시스템
 - 다중 서버의 로컬 디스크를 하나의 디스크처럼 서비스
 - 대용량 스토리지 기반 디자인
- 병렬 파일 시스템
 - 다중 디스크에 쪼개져 있는 하나의 파일에 동시 접근
 - Parallel applications을 위해 디자인
- C 기반 오픈 소스



PVFS2 의 특징

특징	내용	장단점
다양한 파일 시스템 접근 방법	1. API 이용 - libpvfs 2. 커널모듈 이용 (UNIX file system API, POXIS)	1. 응용 프로그램 수정 2. 응용 프로그램 수정 불필요 (속도 떨어짐)
단일 파일 병렬 접근	쪼개진 파일을 동시에 접근하여 취득 가능	1. 접근 속도 빠르다. 2. 처리량 증가
데이터 복제 배치 불가	같은 내용의 데이터 청크가 복제되지 않음	1. Write 시 응답속도가 빠르다. 2. 클러스터 장애 시 데이터 손실 가능 RAID-5등의 하드웨어적 방법 필요
클라이언트 캐시 없음	모든 Write를 IO 노드로 바로 복사	1. 복잡한 cache consistency protocol 불필요 2. 단점: small write가 많으면 성능 저하
단일 파일의 2 Write 허용	단일 파일의 다른 데이터 청크에 동시 Write 가능	1. 빠른 속도 2. 같은 청크의 2 Write 발생할 가능성 있음
확장 된 속성 제공	1. Metadata type에 End-of-File(EoF) offset 이라는 확장된 속성 제공 2. 파일의 분산 정보를 저장	1. File Appending 개발 시 EoF 속성 활용
설치의 편리	5단계 정도의 간단한 설치 과정	1. 단순한 설치과정으로 편리 2. 다양한 옵션이 제공되지 않아 관리의 한계 존재

04. 개발 내용

개발 내용

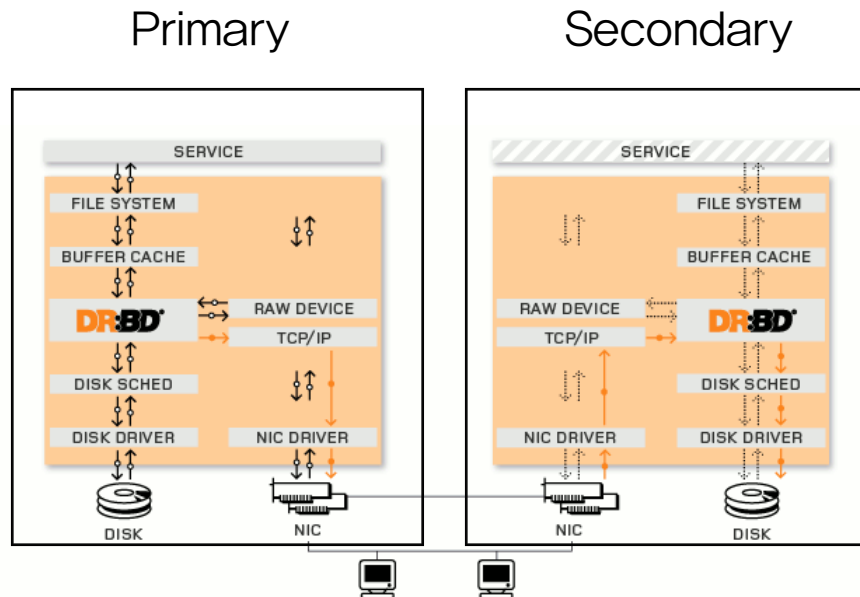
개발 내용

- 1 Metadata Server 이중화
- 2 Scale 보장을 위한 동적 확장 지원
- 3 Small Write 처리 속도 성능 개선
- 4 개발자 가이드 및 API 문서

05. 적용 방법론

Metadata Server 이중화

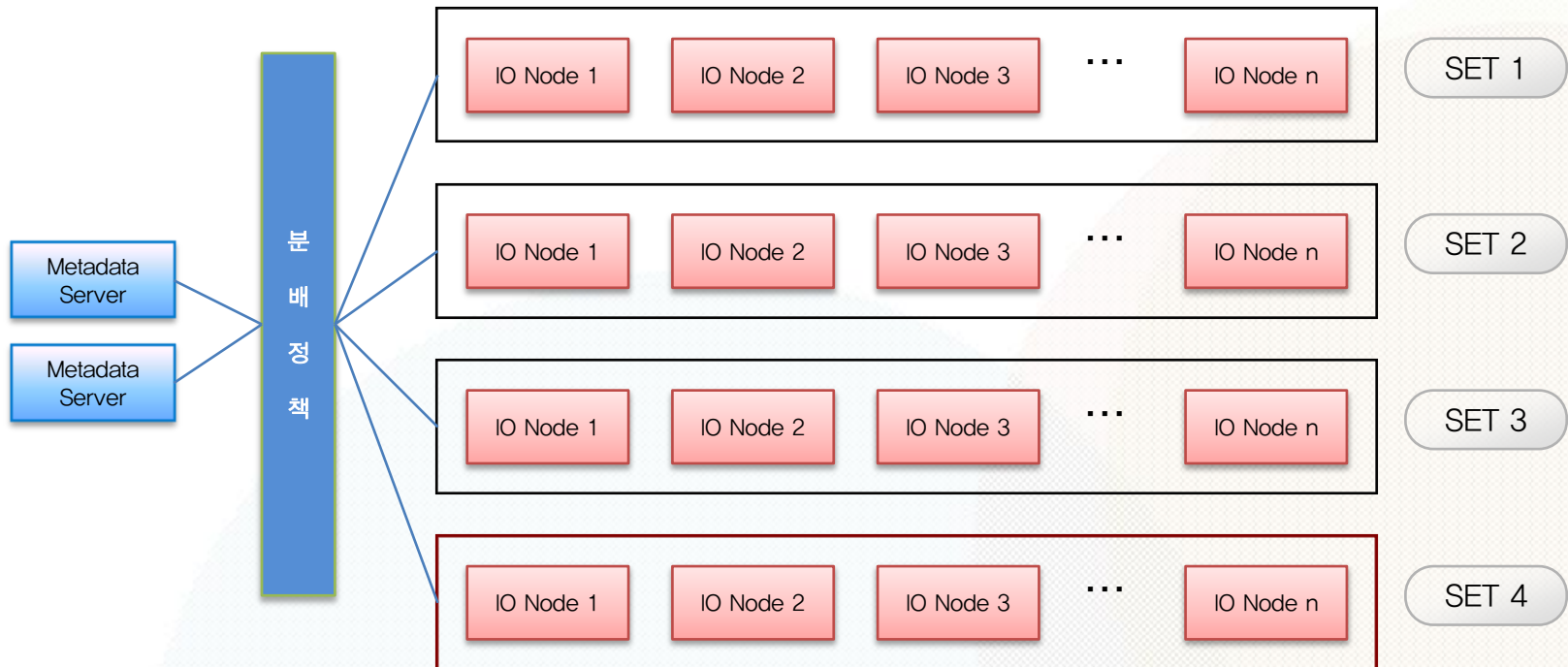
- DRBD (Distributed Replicated Block Device) 를 이용하여 이중화
- Metadata Server 를 Primary, Secondary 로 HA 구성



DRBD를 사용하여 이중화

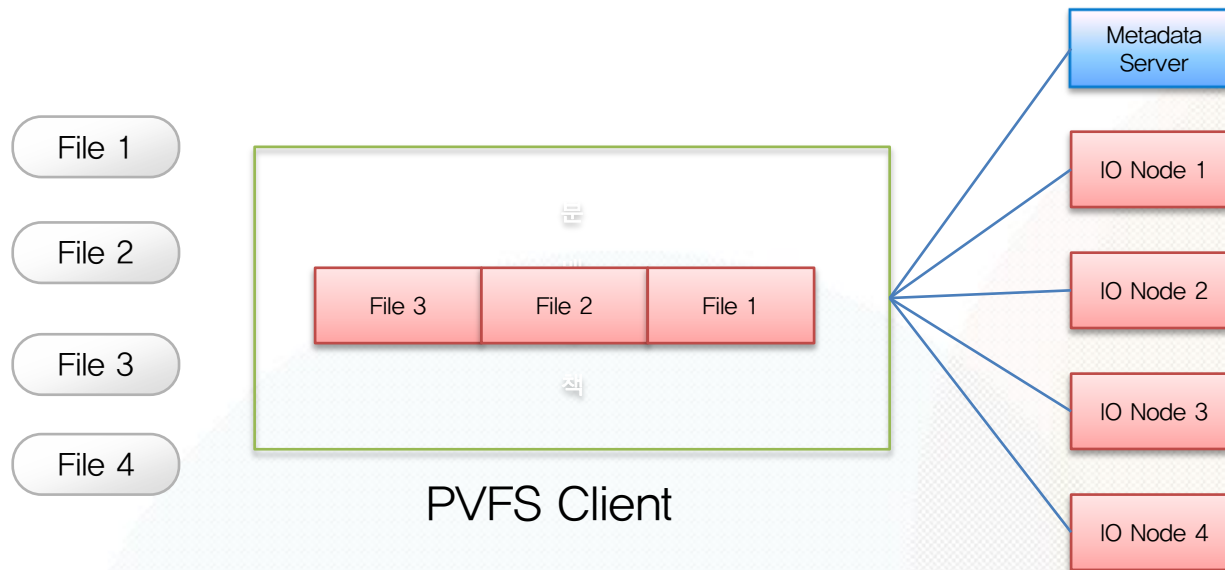
Scale 보장을 위한 동적 확장 지원

- IO Node 를 여러 대를 하나로 묶은 SET 으로 구성
- 동적 확장 시 SET 단위로 IO Node 추가
- PVFS 의 distribution 정책을 수정하여 IO NODE SET 선택



Small Write 처리 속도 성능 개선

- Small Write 처리를 위하여 PVFS Client 수정
- PVFS Client 에 Small Write 에 대한 Memory Caching 추가
- Network 으로 데이터 전송 시 일정 부분 Packing 하여 전송



개발자 가이드 및 API 문서

- PVFS2 설치문서 (Single Host)

<http://www.kldp.net/projects/pvfs2-cloud/forum/316277/707/PVFS2install-asinglehost.pdf>

- PVFS2 설치문서 (VM Cluster)

<http://www.kldp.net/projects/pvfs2-cloud/forum/316286/709/PVFS2-virtualmachinecluster.pdf>

- Amazon EC2 기본 구조 (Load Balancer 등)

<http://www.kldp.net/projects/pvfs2-cloud/forum/316289/710/AWS.pdf>

- PVFS High Availability Clustering using Heartbeat2.0

<http://www.kldp.net/projects/pvfs2-cloud/forum/316341/715/PVFSHighAvailabilityClusteringusingHeartbeat2.0.pdf>

- CloudHarmony에서의 디스크 I/O 와 네트워크 퍼포먼스 벤치마크

<http://www.kldp.net/projects/pvfs2-cloud/forum/316356/717/cloud+harmony+benchmark.pdf>

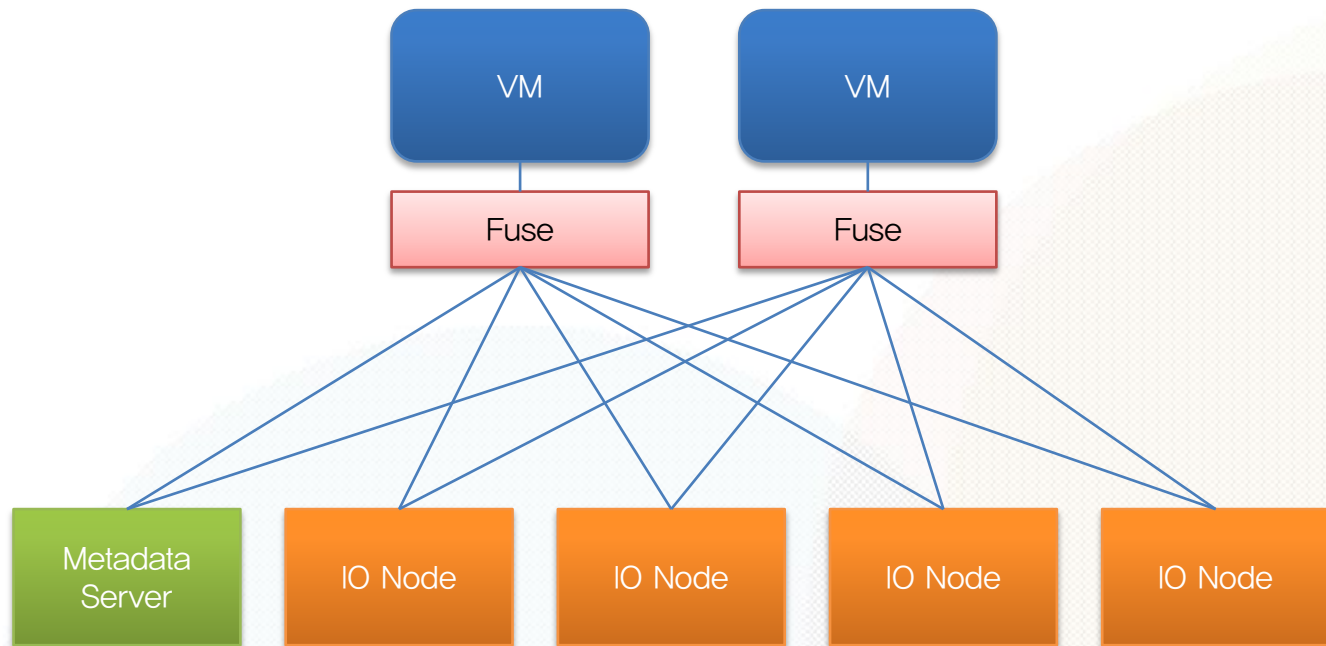
- Xen의 I/O PATH 정리

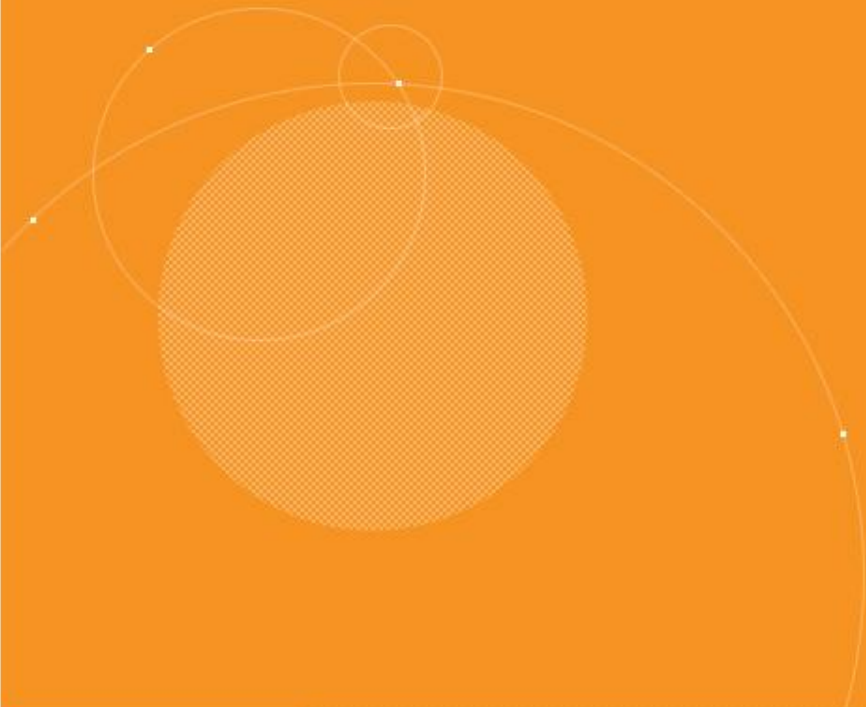
<http://www.kldp.net/projects/pvfs2-cloud/forum/316408/756/Xen-IO.pdf>

06. 서비스 현황

Cloudit File Server 활용

- 1 Metadata Server
- 16 IO Node
- 81T Storage
- 4T Media Data





감사합니다...
(Q&A)