

# OASIS : IP 기반의 대규모 광역 스토리지 시스템

## OASIS : Large-scale, wide-area storage system based on IP

김홍연, 김영철, 진기성, 김영균, 이미영  
한국전자통신연구원

Kim Hong-Yeon, Kim Young-Chul, Jin Ki-Sung,  
Kim Young-Kyun, Lee Mi-Young  
Electronics and Telecommunications  
Research Institute

### 요약

본 논문은 한국 전자통신연구원에서 개발하고 있는 IP 기반의 대규모 광역 스토리지 시스템에 관한 논문이다. OASIS는 IP 네트워크 상의 수백~수천 명의 동시 사용자를 효과적으로 지원하고 또한 네트워크 스토리지 서비스를 WAN 범위까지로 확대 시키는 것을 목적으로 하는 스토리지 시스템이다. 이를 위해 IP 기반 스토리지 연결망 기술과 객체 기반 스토리지 기술 그리고 클러스터 기반 서버 구조 기술을 활용한다. 본 시스템은 특히 FTTH, WiBro 등 차세대 통신망의 도입에 따라 점차 현실성을 얻어가는 IP 기반 스토리지 서비스에서 대규모 스토리지 서비스를 위한 목적으로 활용될 수 있다.

### Abstract

In this paper, a large-scale, wide-area storage system based on IP is proposed which is under development. OASIS is a storage system enforced with very high-scalability up to hundreds and thousands of clients over IP network, and it is able to extend the service to the wide area network. For this purpose, we adopt an storage interconnection technology based on IP, the object based storage technology and a clustered server architecture which provides high-scalability and availability to our system. This system can be utilized to provide network storage service which gains more reality with the incoming FTTH and WiBro services.

## I. 서론

최근 유무선 네트워크의 급격한 속도 향상과 함께 네트워크상에서 스토리지 서비스가 점차 현실성을 얻어가고 있으며 이와 함께 전통적인 네트워크 기반 스토리지 기술인 NAS(network attached storage) [1], SAN(storage area network)[2]에 기반한 서비스와 함께, IP상에서 가상 스토리지 장치 기술을 이용한 SpeedDisk[3], 웹을 기반으로 하는 WebDisk 등의 다양한 기술이 출현하고 있다. 이들은 모두 네

트워크상에서 파일 또는 스토리지를 공유하고자 하는 공통된 목적을 충족시키기 위한 방안으로써 현재 다양한 분야에서 활용 되고 있으나 보다 적극적인 대규모 스토리지 서비스를 위해서는 다음과 같은 한계를 일정부분 가지고 있다.

첫째, 인터넷과 같은 개방된 네트워크상에서 사용할 경우 프로토콜 자체가 가지는 특성에 의한 성능 저하[4]뿐 아니라 심각한 보안 문제[5]에 직면할 수 있다.

둘째, LAN 환경이라 할지라도 서버 자체의 구조적 병목 현상으로 사용자수 증가(수천 이상)에 대한 대응에 한계를 가지고 있다[3,4,5].

셋째, FC에 기반한 SAN의 경우 구축 비용이 크므로 제한된 환경에서만 사용 가능하다.

본 논문에서는 근거리 뿐 아니라 광역망에서 대규모 클라이언트 및 스토리지를 안전하고 효과적으로 지원할 수 있는 기술로서 OASIS(Object storage Architecture for Scalability, Intelligence and Security)를 설계하고 구현한다.

OASIS는 IP 기반 스토리지 연결망 기술인 iSCSI 기술, 현재 표준화가 진행 중인 OSD(object storage device) 기술, 그리고 대규모 스토리지 서버를 위한 OSA(object-based storage architecture) 기술에 기반하며 IP상에서 대규모 스토리지 서비스를 위한 다양한 기능을 제공한다.

본 논문의 구성은 다음과 같다. II장에서는 관련된 구조로서 OSD 기술, OSA 기술에 대하여 설명한다. III장에서는 IP 네트워크상에서 대규모 스토리지 서비스 방법을 제공하는 OASIS의 구조 및 기능에 대하여 설명한다. IV장에서는 시스템의 주요 동작 흐름을 다루고 V장에서 결론을 맺는다.

## II. 관련 연구

OSD는 전통적인 블록 기반 스토리지 장치와 달리 임의 크기의 객체를 기반으로 하는 지능형 스토리지로서 Carnegie Mellon 대학의 Parallel Data Lab.에서 수행한 NASD(network attached secure disk) 프로젝트[5]와 Active Disk 프로젝트[6]에서 개념이 시작되었고, 최근에는 HP, Intel, Seagate, IBM과 같은 기업에서 자사의 스토리지 제품군에 객체 기술을 반영하려는 연구를 진행하고 있다. 초기 OSD 연구에서 얻은 기술적 결과는 스토리지 산업체 표준화 단체인 SNIA(storage networking industry association)에서 OSD Technical Working Group을 통해 기존

SCSI 인터페이스를 기반으로 SCSI OSD 인터페이스 표준[7,8]으로 확장 되고 있다.

OSD를 활용한 가시적인 효과는 파일 시스템의 기능중 하나인 스토리지 공간 관리 기능을 OSD로 이관(offload)하는 것이다[4]. 이를 파일 시스템에서 활용할 경우 파일 시스템은 보다 경량화 될 수 있으며 또한 OSD장치의 추가시 스토리지 공간의 확장과 함께 공간 관리를 위한 메타데이터 처리 능력의 향상 효과를 동시에 기대할 수 있다.

OSA는 바로 이러한 OSD의 특성을 활용한 대규모 스토리지 서버 구조이다[4]. OSA는 개념적으로 클라이언트, 메타데이터 서버 그리고 객체 기반 스토리지 서버들이 IP기반의 네트워크로 상호 연결된 구조를 가지며, 메타데이터 서버는 파일 시스템 메타데이터를 관리하고, 공간 관리를 위한 메타데이터 처리 및 자료 저장 기능을 객체 기반 스토리지 클러스터가 부담하는 구조이다.

## III. OASIS 시스템

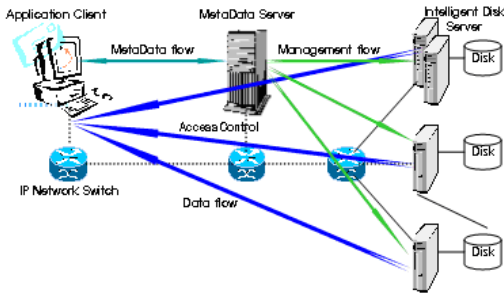
### 1. 시스템 목표

OASIS는 다수의 사용자들이 이기종 플랫폼 환경에서도 데이터를 안정적으로 공유하여, 언제/어디서나 데이터 접근 서비스를 제공받을 수 있는 파일 시스템 S/W 및 스토리지 H/W 플랫폼 개발을 목표로 하며 주요 목표는 다음과 같다.

- IP 기반의 개방된 유무선망을 통한 일관되고 안전한 스토리지 접근
- 대규모 클라이언트 및 스토리지에 적합한 서버 구조

### 2. 시스템 구조

OASIS는 그림 1과 같이 OSA를 기반으로 하며, 각 구성요소의 역할은 다음과 같다.



▶▶ 그림 1. OASIS 시스템 구조

- 클라이언트: 파일 입출력을 수행하는 클라이언트가 스토리지를 사용할 수 있도록 파일 시스템을 연결하여, 파일 입출력에 필요한 표준의 파일 입출력 인터페이스를 처리하는 기능을 수행한다.
- 메타데이터 서버: 파일 시스템이 운영되는데 필요한 fileset 및 이름 공간 메타데이터를 관리하는 기능, 시스템의 가용성을 높이는 클러스터링 구성 관리 기능 등을 제공한다.
- 객체 기반 스토리지: 객체기반 스토리지 장치(object-based storage device: OSD)에 저장될 수 있는 논리적인 객체 저장 메타데이터 관리와 객체 입출력을 수행한다.

각각은 독립된 컴퓨팅 노드(즉, 서버)에서 동작되고, 이들은 Gigabit 네트워크를 통해 서로 연결되며, 메타데이터 서비스 프로토콜 및 iSCSI/OSD 프로토콜을 이용하여 서로 교신한다.

### 3. 시스템 기능

OASIS 시스템이 제공하는 주요 기능은 시스템 관리, 파일 시스템 구성, 파일 입출력 처리, 가용성 보장, 확장성 지원, 고속 파일 입출력 처리, 보안성 지원 등으로 분류하며, 각 분야별 세부 기능은 다음과 같다.

#### 3.1. 시스템 관리

OASIS 시스템 관리에는 OASIS 시스템 기동 및

종료, OASIS 시스템에서 관리하는 OSD들의 구성 및 관리, 메타데이터 서비스의 가용성을 위한 클러스터링 구성 및 관리, 그리고 시스템 고장시 복구 관리 등이 포함된다. 이러한 시스템 관리 기능들은 OASIS 시스템의 여러 관리 도구(utility) 형태로써 제공된다.

#### 3.2. 파일 시스템 구성

OASIS 시스템에서 제공하는 파일 시스템(OASIS FS라 칭함)은 fileset 개념으로 관리된다. fileset은 디렉토리와 파일들로 구성되는 논리적인 가상의 볼륨이며, 클라이언트는 OASIS/FS에 하나 이상의 fileset을 생성하고 이들 단위로 mount/umount를 수행한다.

#### 3.3. 파일 접근 인터페이스 지원 및 처리

OASIS 시스템은 OASIS FS를 통해 클라이언트 응용이 네트워크에 연결된 스토리지에 존재하는 파일(또는 디렉토리)에 대한 접근 기능(생성, 파일 읽기/쓰기, 변경, 파일 개방, 링크 등)을 제공한다. 파일 접근 연산은 응용이 동작하는 클라이언트 운영 플랫폼(운영체제)에 따라 처리하는 인터페이스가 상이하며, Linux 플랫폼 환경에서 동작하는 OASIS FS는 POSIX 표준 인터페이스에 준하는 기능을 제공한다.

#### 3.4. 가용성 보장

OASIS 시스템은 파일 데이터를 저장하는 스토리지 장치와 파일 시스템 메타데이터를 관리하는 메타데이터 서버 노드에 장애가 발생한 경우에도 클라이언트에 지속적인 파일 접근 서비스를 제공하는 가용성 보장 기능을 제공한다. 가용성 보장을 위해서 아래와 같이 2가지 종류의 세부 기능들을 제공한다.

- 파일 데이터 가용성: 파일 데이터는 스토리지 클러스터를 구성하는 다수의 OSD들에 분산 저장(striping)하고 특정 OSD 노드에 장애가 발생한 경우를 대비하여 OSD간의 RAID 기능(RAID

level 5)을 제공한다.

- 메타데이터 가용성: 메타데이터는 메타데이터 서버 클러스터를 구성하는 다수의 메타데이터 서버 노드에 분산 저장하고 장애가 발생하였을 경우 이를 정상 서버 노드가 백업하는 기능을 제공한다.

### 3.5. 확장성 지원

OASIS 시스템은 보다 많은 클라이언트들에 파일 서비스를 지원하기 위하여 메타데이터 관리 블록의 클러스터링 환경에서 부하 분산(load balance) 기능을 제공한다. 또한 시스템 운영 중에 추가적인 스토리지 공간 확장을 수행하는 기능을 제공한다.

- 부하 분산: 메타데이터는 클러스터링에 의해 분산 관리하여, 클라이언트의 메타데이터 서비스 요청이 특정 노드에 집중되지 않도록 부하를 분산한다.
- OSD 확장: OASIS 시스템에서 관리하는 스토리지에 대한 공간 확장을 위해 OSD 클러스터에 새로운 OSD 장치의 추가를 지원한다.

### 3.6. 고속 파일 입출력 처리

OASIS 시스템은 클라이언트에서 요청된 파일 접근 연산을 고속으로 처리할 수 있도록 파일 관리 블록에서 메타데이터와 데이터에 대한 캐칭(caching) 기능과 네트워크로 연결된 다수의 OSD로부터 객체 데이터를 병렬로 입출력 처리하는 기능을 제공한다.

### 3.7. 시스템 보안

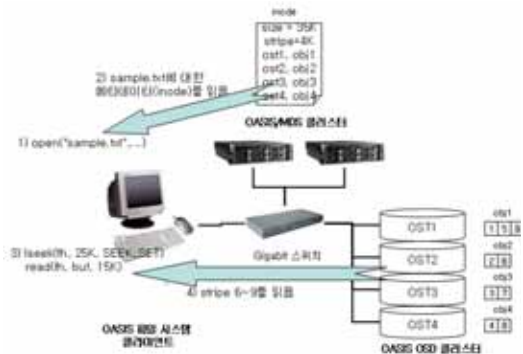
OASIS 시스템은 시스템의 안전성을 제고하기 위하여 파일 시스템의 고유 기능인 접근 제어 처리 기능과 네트워크 환경에서 필수적인 사용자 인증, 네트워크 전송 데이터 보호 등과 같은 보안 처리 기능을 제공한다.

- GSSAPI 기반의 개방형 보안 모델 지원

- Capability 기반의 클라이언트와 스토리지간 인증
- Posix ACL 기반의 접근 제어

## IV. 동작 흐름

OASIS의 주요 기능 중 파일 읽기 기능의 동작 흐름은 그림2와 같다. 먼저 클라이언트에서 메타데이터 서버에게 접근하려는 파일에 대한 메타데이터인 Inode를 요청한다. Inode 속에는 해당 파일의 실제 자료를 저장하고 있는 객체들의 리스트를 가지고 있으며 이를 이용하여 읽고자 하는 위치에 대응되는 객체 식별자를 구하여 최종적으로 원하는 자료를 해당 객체 저장 장치로부터 읽는다.



▶▶ 그림 2. 파일 읽기 동작 흐름

## V. 결론

네트워크 기술의 발전으로 인해 네트워크 기반 스토리지 기술은 점차 그 활용 범위를 넓히고 있다.

본 논문에서는 IP 네트워크 환경 하에서 대규모 스토리지 서비스를 효과적으로 지원하기 위한 시스템으로서 OSA에 기반한 OASIS의 구조 및 기능에 대해 소개하였다.

제안된 시스템은 대규모 스토리지 서비스에서 NAS등의 확장성 문제, SAN의 비용 문제 등을 극복

하기 위해 활용될 수 있으며 FTTH, WiBro 등의 서비스 출현과 함께 점차 현실성을 얻어갈 것으로 예측되는 IP 기반 스토리지 서비스에서 보다 효율적인 스토리지 서버로 활용될 수 있다.

#### ■ 참고문헌 ■

- [1] Garth A. Gibson, et al, "Network Attached Storage Architecture," Communications of the ACM, Vol.43, No.11, 2000.
- [2] White paper, "Introduction to Storage Area Networks", IBM, <http://www.ibm.com>
- [3] 김경배 "블록 전송 기법을 이용한 IP기반 스토리지 시스템:SpeedDisk", 한국정보과학회추계논문집, 2003
- [4] Intel Corporation, "Object-Based Storage: The Next Wave of Storage Technology and Devices", White paper, [ftp://download.intel.com/labs/storage/download/odswhitepaper\\_0703.pdf](ftp://download.intel.com/labs/storage/download/odswhitepaper_0703.pdf), July 2003
- [5] NASD: Network Attached Secure Disk, Carnegie Mellon univ., <http://www.pdl.cmu.edu/NASD/index.html>
- [6] <http://www.pdl.cmu.edu/Active>
- [7] T10 Organization, "Information technology SCSI Object-Based Storage Device Commands(OSD)", <http://www.t10.org/ftp/t10/drafts/osd/osd-r09.pdf>, 19 Feb. 2004.
- [8] <http://www.snia.org>