

대용량 Abyss Storage의 KOREN 네트워크 기반 국내 및 해외 실증 테스트

(KOREN based Domestic and International Verification Test of Mass Abyss Storage)

차병래^{***}, 차윤석^{**}, 최명수^{**}, 박선^{*}, 김종원^{*}

(ByungRae Cha, YoonSeok Cha, MyeongSoo Choi, Sun Park, and JongWon Kim)

요약

ICT 분야의 메가 트렌드는 IoT, Bigdata, 그리고 Cloud Computing 으로 집약된다. 이러한 메가 트렌드들은 독립적으로 운영되지 않으며, 이를 지원하기 위한 백그라운드에는 대용량 컴퓨팅 기술이 필요하듯 대용량 스토리지 기술이 필수적이다. 오픈소스 Ceph 기반의 대용량 스토리지의 성능 평가를 위하여 교육망인 KOREN 네트워크를 이용한 국내 및 해외 사이트와의 Abyss Storage의 실증 테스트를 수행한다. 또한, 스토리지 자체의 성능 평가를 위하여 스토리지 매체와 네트워크 본딩의 성능 테스트를 수행한다. 스토리지 매체별로 실질적인 속도 차이는 크지만 수행된 스토리지 매체 테스트에서는 크게 성능 차이를 보이지 않았다. 이를 해결하기 위한 방안으로 네트워크 가속화를 통한 성능 향상을 얻을 수 있었다. 또한 KOREN 망을 이용한 국내 및 해외 사이트를 연결하여 Abyss Storage의 내부 및 외부 네트워크의 실질적인 성능 테스트를 수행하였다.

■ 중심어 : 대용량 스토리지; Abyss Storage; 코넷; 네트워크 본딩

Abstract

The trends in ICT are concentrated in IoT, Bigdata, and Cloud Computing. These mega-trends do not operate independently, and mass storage technology is essential as large computing technology is needed in the background to support them. In order to evaluate the performance of high-capacity storage based on open source Ceph, we carry out the demonstration test of Abyss Storage with domestic and overseas sites using educational network KOREN. In addition, storage media and network bonding are tested to evaluate the performance of the storage itself. Although there is a substantial difference in aspect of the physical speed among storage medias, there is no significant performance difference in the storage media test performed. As a solution to this problem, we could get performance improvement through network acceleration. In addition, we conducted actual performance test of Abyss Storage internal and external network by connecting domestic and overseas sites using KOREN network.

■ keywords : Mass Storage; Abyss Storage; KOREN; Network Bonding

I. 서론

대부분의 신기술은 제품의 성능을 향상시키며, 이러한 기술을 존속적 기술이라 하며, 존속적 기술은 각각의 성격상 단절적이거나 급진적일 수도 있지만 점진적인 성격을 갖고 있는 것들도 많다. 모든 존속적 기술이 공통적으로 갖고 있는 특징은 주요 시장에서 활동하는 주류 고객들이 기대하는 수준에 맞추어 기존 제품의 성

능이 향상된다는 것이다. 그러나 때로 파괴적인 기술이 생겨날 때가 있으며, 혁신적 기술은 과거에 통용됐던 것과 아주 다른 가치명제(value proposition)를 시장에 선보이게 된다. 일반적으로 혁신적 기술은 주류 시장에서 기존 제품들보다 성능이 떨어지며, 혁신적 기술에 기초한 제품들은 일반적으로 더 싸고, 더 단순하고, 더 작고, 사용하기 더 편리하다. 따라서 [그림 1]의 예시와 같이 기존의 선도 기업들의 존속적 스토리지 기술에서 신생의 중소기업들을 위한 혁신적인 스토리지 기술로 전환 및 성장의 기회를 얻기 위해

* 정회원, 광주과학기술원 전기전자컴퓨터공학부

** 정회원, 제노테크(주)

이 논문은 2016년도 중소기업청 창업성장 기술개발사업(오픈소스 기반의 대용량 분산 Abyss Storage와 네트워킹 가속화 기술 개발)의 결과물임 (No. S2385742)

접수일자 : 2017년 03월 06일

수정일자 : 2017년 03월 24일

게재확정일 : 2017년 03월 28일

교신저자 : 김종원 e-mail : jongwon@gist.ac.kr

오픈소스 기반의 대용량 분산 스토리지 및 네트워크 가속화 기술 개발을 하려고 하며, 그 중에 대용량 분산 스토리지 개발을 위해 오픈소스 Ceph를 사용하고 있다.

Ceph는 인텔 프로세서 기반 범용 하드웨어를 사용해 대규모 소프트웨어 정의 스토리지를 구현하기 위해 시작된 오픈 소스 프로젝트이다. 따라서 비용 및 제한, 그리고 독점 없이 스토리지 서비스가 가능하다. 이것은 통합, 분산, 높은 확장성 및 신뢰성을 제공하는데 이는 현재와 미래의 비정형 데이터에 반드시 필요하며, 세계적으로 대용량 스토리지 요구가 확장하고 있다. 그렇기에 우리는 데이터 신뢰성 및 성능에 영향을 주지 않는 Multi-Exabyte 레벨의 확장성 있는 스토리지 시스템이 필요한 상황이다. 본 논문에서는 오픈소스 Ceph 기반의 대용량 스토리지의 성능향상을 위해 디스크 매체별로 성능 테스트와 네트워크 분당, 그리고 KOREN 네트워크를 이용한 국내외의 네트워크 트래픽 테스트를 수행하였다.

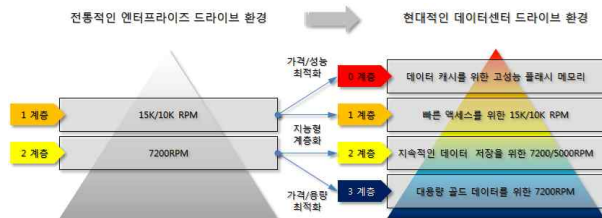


그림 1. 존속적 스토리지 H/W 기술에서 혁신적 스토리지 H/W 기술의 세분화 및 지능화

II. 관련 연구

스토리지 기술은 크게 DAS(Directed Attached Storage), NAS(Network Attached Storage), 그리고 SAN(Storage Area Network)으로 분류되며, [그림 2]와 [그림 3]과 같이 나타낼 수 있다[1].

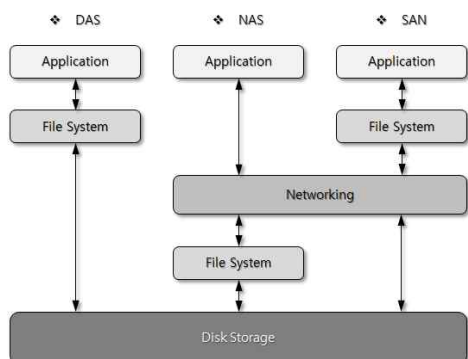


그림 2. DAS, NAS, 그리고 SAN의 개념

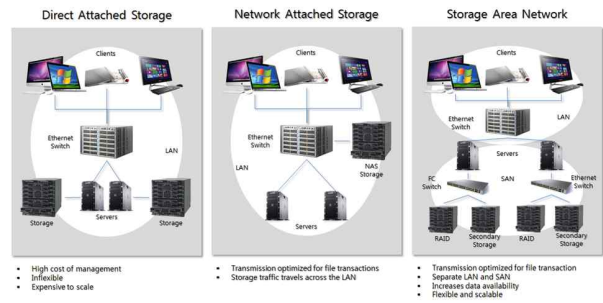


그림 3. 스토리지의 분류 및 특징

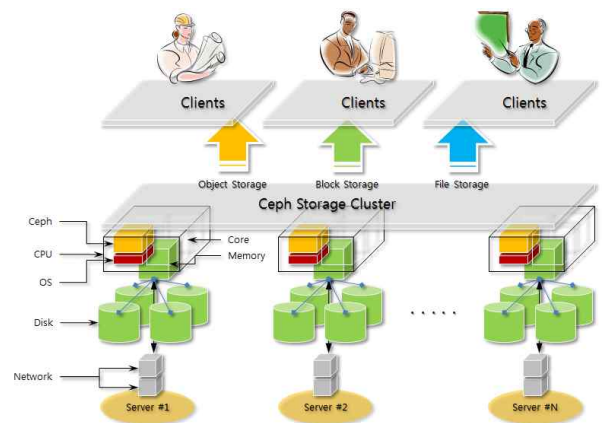


그림 4. Ceph 스토리지 클러스터의 개념도

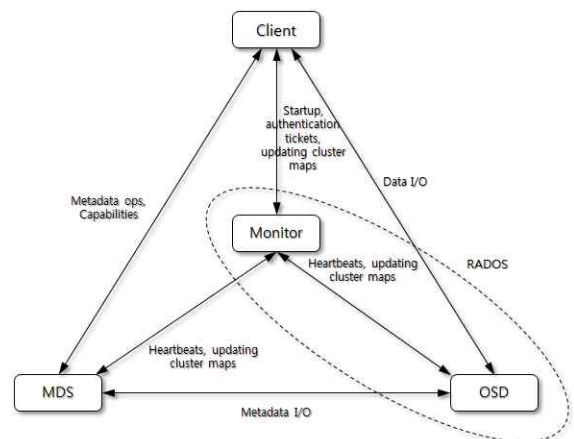


그림 5. Ceph 시스템의 아키텍처와 상호 작용

대용량 스토리지를 위한 오픈 소스 Ceph는 페타 바이트 규모의 Linux 분산 파일 시스템이며, UCSC(University of California, Santa Cruz)에서 Sage Weil의 스토리지 시스템 관련 박사과정 연구 프로젝트로 시작되었으며, [그림 4]는 Ceph 스토리지 클러스터의 개념도를 나타낸 것이며, 오픈소스 Ceph의 목표는 다음과 같이 정의된다.

- 용량을 페타 바이트 수준으로 손쉽게 확장
- 가변적인 워크로드를 효과적으로 처리할 수 있는 고성능 및 대역폭
- 강력한 신뢰성

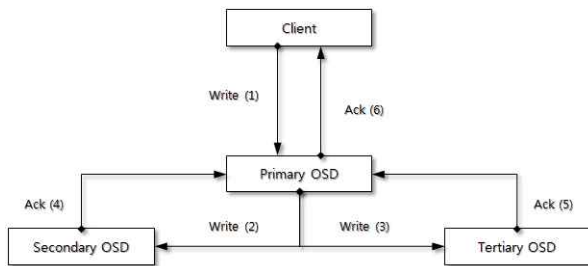


그림 6. Ceph의 Self-Healing과 3중 복사 절차

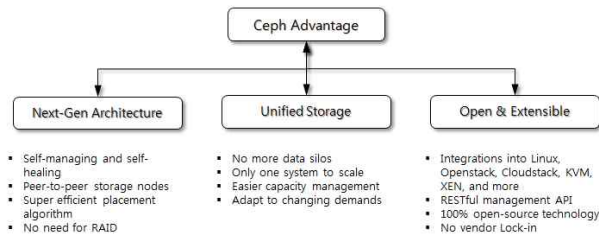


그림 7. 오픈 소스 Ceph의 장점

오픈소스 Ceph 시스템 아키텍처와 상호작용을 [그림 5]와 같이 표현할 수 있으며, Ceph 시스템은 크게 4 가지 세그먼트로 분류할 수 있다.

- 데이터의 사용자인 클라이언트
- 분산 메타데이터를 캐시 및 동기화하는 메타데이터 서버
- 데이터와 메타데이터를 오브젝트로 저장하고 기타 주요 기능을 구현하는 오브젝트 스토리지 클러스터
- 모니터링 기능을 구현하는 클러스터 모니터

또한 오픈소스 Ceph의 Self-Healing과 데이터의 3중 복사 절차를 [그림 6]에 간략하게 나타낸다. 오픈 소스 Ceph의 장점으로는 차세대 아키텍처, 통합형 스토리지, 그리고 개방성 및 확장성이며, [그림 7]과 같이 나타낼 수 있다[2]. 오픈 소스 Ceph은 풀 수준에서 Ceph클러스터의 성능을 측정하는 데 사용할 수 있는 RADOS bench라고 하는 내장된 벤치마킹용 SW 도구가 함께 제공된다[3].

III. 대용량 Abyss Storage의 설계 및 구현

SMB를 위한 대용량 Abyss Storage의 서버 구성은 [그림 8]과 같이 나타낼 수 있으며, 스토리지 서버의 H/W 사양은 [표 1]에 정리하였다. 또한 실제적으로 Abyss Storage 제품의 H/W 프로토타입 개발과 제품 양산을 위한 제품 케이스의 3D 렌더링 이미지, 그리고 3D 프린팅 결과물을 [그림 9]와 같이 나타냈다[4, 5].

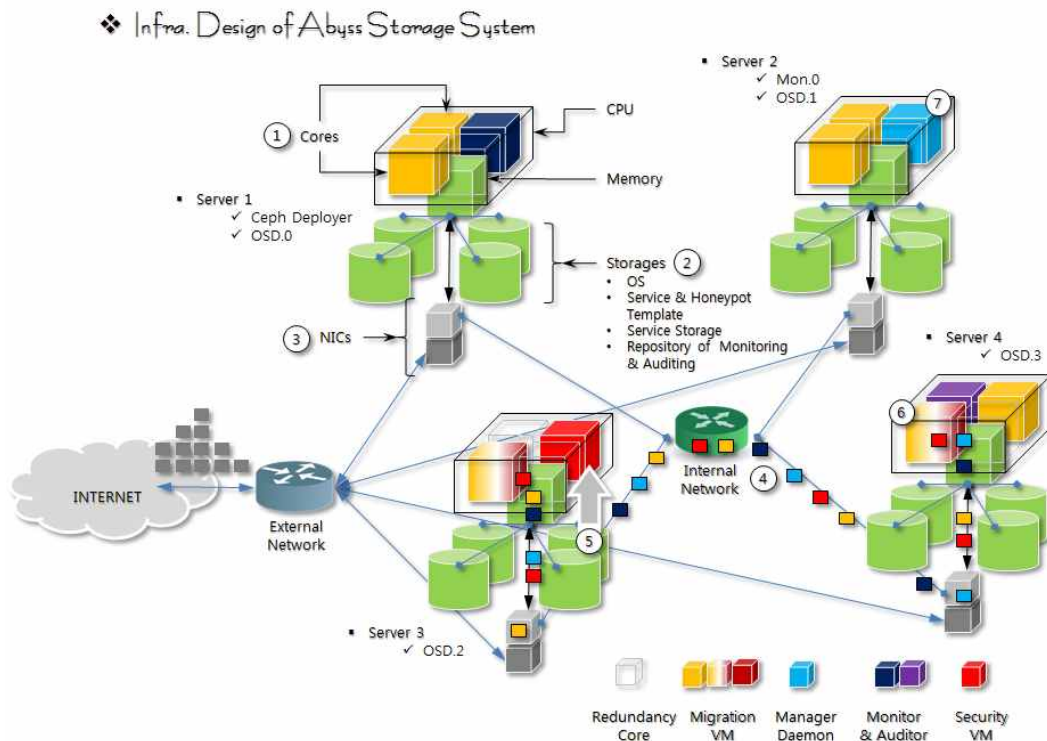


그림 8. Abyss Storage 구성을 위한 개념도

표 1. Abyss Storage 서버의 사양

항목	사양
CPU	Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz
MEM	32GB
NET Device	Intel Corporation I211 Gigabit Network Connection / Intel Corporation Ethernet Connection (2) I218-V
OS	Ubuntu 16.04 LTS Desktop

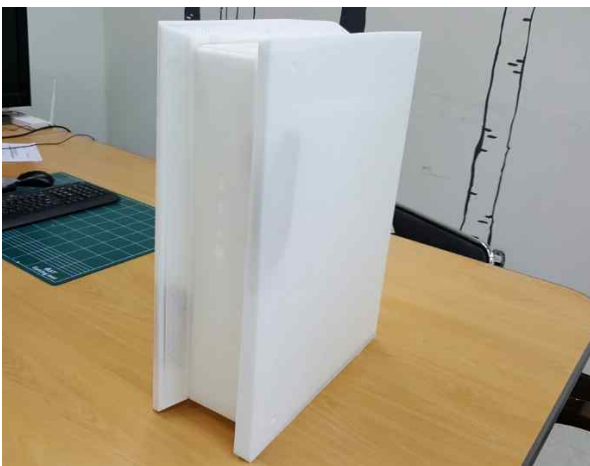
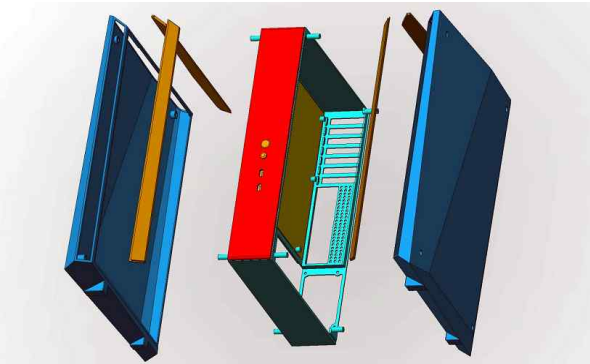


그림 9. Abyss Storage의 H/W 프로토타입과 제품 케이스의 3D 랜더링 이미지, 그리고 3D 프린팅 결과물

IV. Abyss Storage의 성능 테스트

Abyss Storage의 성능 향상을 위하여 스토리지의 디스크 매체별 성능 테스트와 스토리지의 내부 네트워크의 가속화를 위한 본딩, 그리고 KOREN 네트워크를 이용한 국내외 네트워크 트래픽 테스트를 수행하였다.

1. 디스크 매체별 성능 테스트

Ceph RADOS bench 테스트를 위한 대용량 Abyss Storage의 서버의 디스크 매체의 타입은 HDD, SSHD(Solid State Hybrid Drive), SSD의 3 가지이며 사양은 [표 2]와 같이 정리하였다.

표 2. Abyss 스토리지의 디스크 타입 및 사양

종류	사양	용량
HDD	• Seagate Barracuda ST2000DM001 (SATA3/7200/64M)	2TB
SSHD	• Seagate Desktop SSHD ST2000DX001 (SATA3/8G/64M)	
SSD	• 삼성전자 850 EVO Series	

Abyss Storage의 디스크 매체들(HDD, SSHD, SSD)을 RADOS Bench SW를 이용하여 10초 동안 쓰기(연속, 무작위)를 10번을 수행하여 평균 IOPS(Input/output Operations Per Second)를 기록하였으며, [그림 10]과 같은 성능을 보였다[6].

Abyss Storage의 디스크 매체에 따른 IOPS 비교

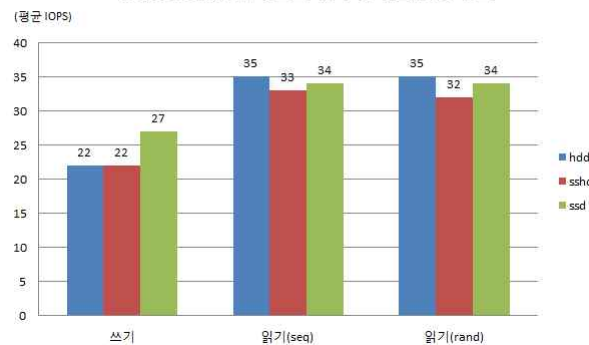


그림 10. 디스크 매체에 따른 IOPS 성능 비교

대용량 스토리지를 구성하는 디스크 매체에 따른 IOPS의 차이는 [그림 10]과 같이 매우 미미했으며, 그 이유는 바로 네트워크 속도가 1GB 이었기 때문이다. 본질적으로 디스크 SSD와 HDD의

성능 차이는 매우 크다. 그러나 위와 같은 결과가 나온 이유는 네트워크 속도가 SSD의 성능을 끌어내지 못했기 때문이며, 향후에 추가적으로 광 랜을 이용한 디스크 매체 및 네트워크 성능 테스트를 수행할 계획이다.

2. 네트워크 가속화를 위한 본딩

Abyss Storage의 네트워크 가속화를 위해서 다음의 [그림 11]과 같이 내부 네트워크를 이중으로 구성하기 위하여 1 GB의 Switch 2개를 VLAN으로 연결한 뒤 각각 스위치 1에서는 서버의 네트워크 카드1, 스위치 2에서는 서버의 네트워크 카드 2에 연결하였다. 서버에서는 네트워크 카드를 하나로 묶어주는 Bonding을 사용하여 하나의 네트워크처럼 설정하였다.

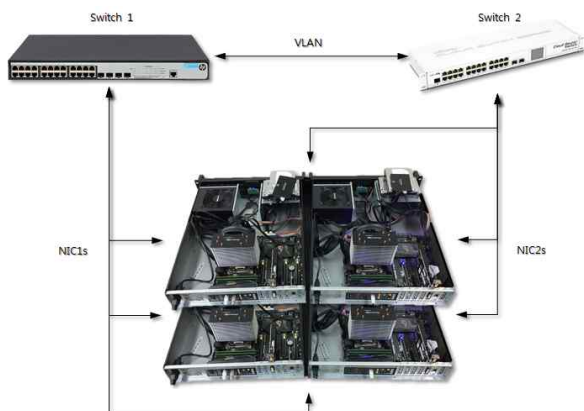


그림 11. Abyss Storage의 네트워크 가속화를 위한 네트워크 본딩 구성

네트워크 가속화 테스트를 위하여 각각 10초 동안 쓰기, 읽기 (Seq), 읽기(Rand) 테스트를 10번 수행하였으며 평균 IOPS의 값을 비교하였으며, 비교 결과는 다음의 [그림 12], [그림 13], 그리고 [그림 14]에 나타났다. 기존 시스템과 네트워크 가속화를 위한 본딩된 시스템의 테스트 결과가 2 배까지는 아니지만 최소 150%이상 큰 폭으로 향상된 것을 확인할 수 있었다.

네트워크에 따른 평균 IOPS(Write)

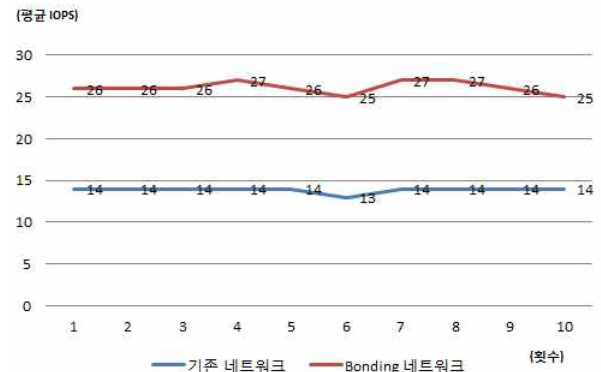


그림 12. 네트워크 본딩에 의한 Write의 IOPS 비교

네트워크에 따른 평균 IOPS(Read_seq)

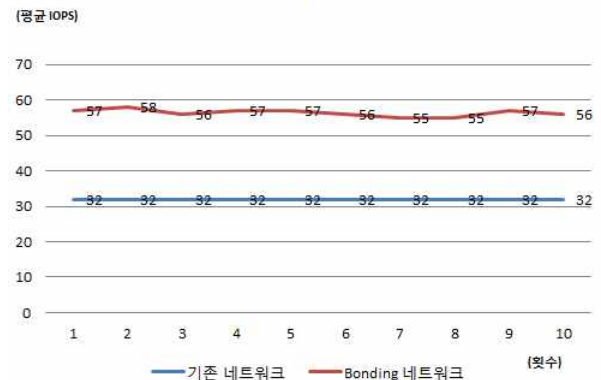


그림 13. 네트워크 본딩에 의한 Read(Seq)의 IOPS 비교

네트워크에 따른 평균 IOPS(Read_rand)

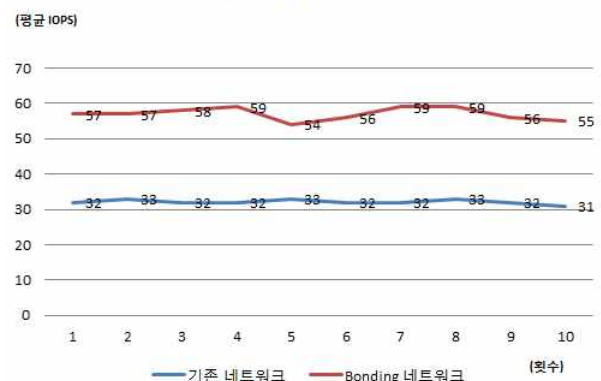


그림 14. 네트워크 본딩에 의한 Read(Rand)의 IOPS 비교

3. 국내외 사이트의 네트워크 테스트

국내의 사이트 GIST와 해외 사이트인 말레이시아의 Myren과의 멀티미디어 데이터의 업로드와 다운로드를 통한 네트워크 트래픽을 테스트를 수행하였다. [그림 15]는 국내외의 실환경 성능 테스트를 위한 테스트베드의 구성도를 나타낸 것이다.

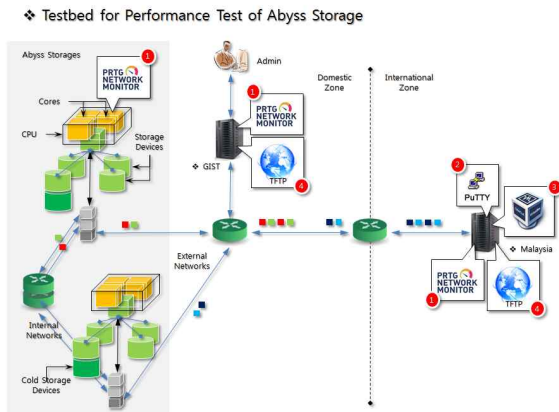


그림 15. 실환경 성능 테스트를 위한 테스트베드 구성도

국내의 GIST 사이트와 개발된 Abyss Storage 간의 파일 용량별 Download 와 Upload 속도를 측정하여 [그림 16]과 [그림 17]에 나타냈다. 더불어 네트워크 실시간 모니터링 도구인 SpeedoMeter[7]를 이용하여 Abyss Storage의 외부 네트워크 트래픽과 Abyss Storage의 클러스터 내부의 네트워크 트래픽을 [그림 18]과 [그림 19]와 같이 측정하였다. 또한, 해외 사이트인 말레이시아의 Myren과 개발된 Abyss Storage 간의 파일 용량별 Download와 Upload 속도를 측정하여 [그림 20]과 [그림 21]에 나타냈다.

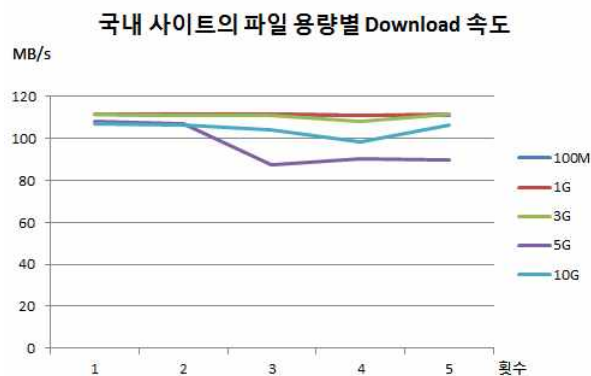


그림 16. GIST와 Abyss Storage의 Download 속도 측정

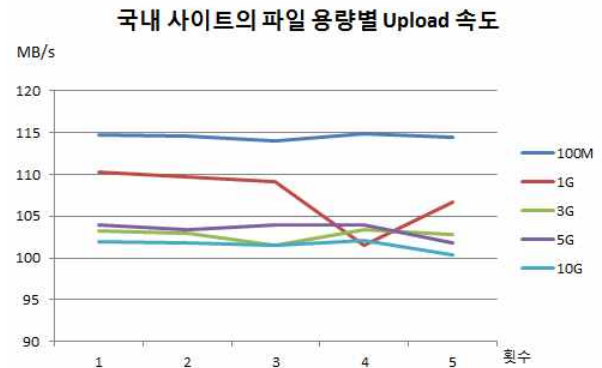


그림 17. GIST와 Abyss Storage의 Upload 속도 측정



그림 18. SpeedoMeter에 의한 외부 트래픽 측정

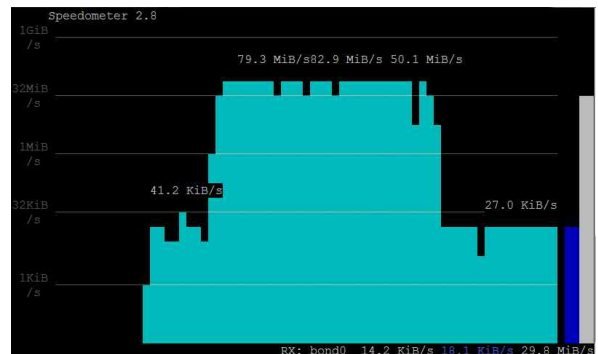


그림 19. SpeedoMeter에 의한 클러스터 내부의 트래픽 측정

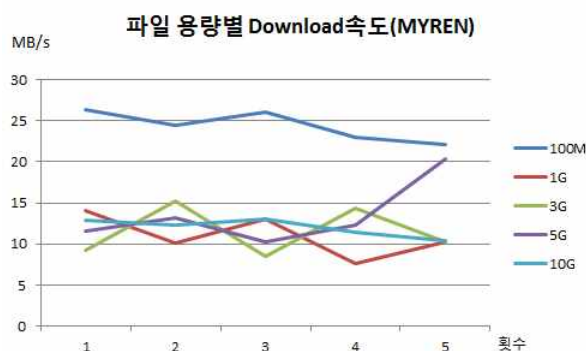


그림 20. 말레이시아 사이트와 Abyss Storage의 Download 속도 측정

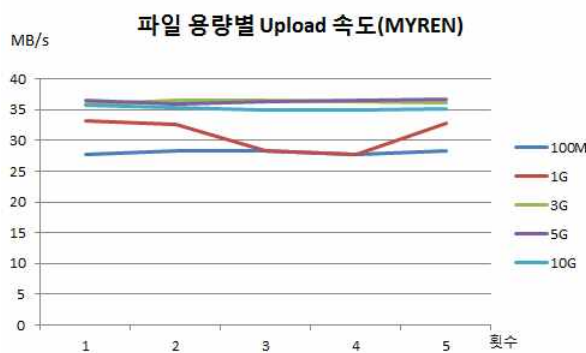


그림 21. 말레이시아 사이트와 Abyss Storage의 Upload 속도 측정

V. 결 론

개발된 대용량 Abyss Storage 프로토타입의 성능 테스트와 KOREN 네트워크를 이용한 국내외의 실환경 테스트를 진행하였으며, 이를 기반으로 성능 향상을 위한 방안을 모색하고자 한다.

성능 향상을 위한 세부적인 테스트로는 스토리지의 디스크 매체별 Read와 Write 성능 테스트, 네트워크 본딩에 의한 테스트, 그리고 국내외 사이트들 간의 네트워크 트래픽 테스트를 진행하였다. 디스크 매체에 따른 IOPS의 차이는 매우 미미했으며, 그 이유는 바로 네트워크 속도가 기가비트(1G)였기 때문이었다. 본질적으로 디스크 SSD와 HDD의 성능 차이는 매우 크다. 그러나 위와 같은 결과가 나온 이유는 네트워크 속도가 SSD의 성능을 끌어내지 못했기 때문이며, 향후에 10G 네트워크 환경에서 추가적인 테스트 및 Software-Defined RAID[8, 9, 10] 기술을 적용할 계획이다. 또한 네트워크 가속화를 위한 네트워크 본딩에 의한 실제적인 성능 향상을 확인하였으며, 국내외 사이트 간의 원활한 네트워크 테스트를 진행하였다.

References

- [1] SAN, NAS, and DAS, "What is the difference between SAN, NAS and DAS?," <http://serverfault.com/questions/81723/what-is-the-difference-between-san-nas-and-das>
- [2] Karan Singh, *Learning Ceph*, PacktPublishing, Jan. 2015.
- [3] Karan Singh, *Ceph Cookbook*, PacktPublishing, pp.270-272, Feb. 2016.
- [4] 차병래, 김지수, 박선, 김종원, "Draft of Mass Distributed Storage System based on PC using Ceph," *SMA 2014*, 2014.
- [5] 차병래, 김용일, "Prototype Design of Mass Distributed Storage System based on PC using Ceph for SMB," *스마트미디어저널*, 제4권, 제3호, 62-67쪽, 2015년 9월
- [6] 차윤석, 조원오, 오수희, 최명수, 차병래, "Abyss Storage의 Disk 타입에 의한 Ceph RADOS의 Benchmarking," *2017 한국통신학회 동계학술대회*, 2017년 1월
- [7] SpeedoMeter, <https://hub.docker.com/r/opennsm/speedometer/>
- [8] 차병래, 박선, 김종원, "분산 저장환경의 데이터공유 및 관리를 위한 소프트웨어 정의 저장 방법," *한국정보통신학회 종합학술대회 논문집 춘계*, 제18권, 제1호, 2014년 5월 30일.
- [9] 차병래, 최명수, 박선, 김종원, "Software-Defined RAID 기반 장애복구 기법과 실증 테스트," *스마트미디어저널*, 제5권, 제1호, 2016년 3월
- [10] 차병래, 박선, 김종원, "Design and Verification of Software-Defined RAID for Hybrid Cloud Storage," *ICCCRI 2016*, 2016년 5월

 저 자 소 개

차병래



2004년 목포대학교 대학원 컴퓨터공학과 졸업(공학박사)
 2005년 호남대학교 컴퓨터공학과 전임 강사
 2009년 ~ 현재 광주과학기술원 전기전자컴퓨터공학부 연구조교수

2012년 ~ 현재 제노테크(주) 대표

<주관심분야: 정보보안, IDS, Neural Network, Cloud Computing, VoIP, NFC, 대용량 스토리지 기술 등>

차윤석



2014년 고려대학교 컴퓨터정보학과 학사

2015년 ~ 현재 제노테크(주) 기업부설 연구소 연구원

<주관심분야: IoT, BigData, Cloud Computing, VoIP, NFC, 대용량 스토리지 기술 등>

리지 기술 등>

최명수



2009년 목포대학교 전자공학과 공학박사

2009년 목포대학교 해양텔레매틱스기술개발센터 박사후연구원

2010년 목포대학교 정보산업연구소 연구전임교수

2015년 ~ 현재 제노테크(주) 기업부설연구소 연구소장

<주관심분야: IoT, Neural Network, Cloud Computing, VoIP, NFC>

박 선



2007년 인하대학교 컴퓨터정보공학과 공학박사

2008년 호남대학교 컴퓨터공학과 전임 강사

2010년 전북대학교 인력양성사업단 박사후 과정

2010년 목포대학교 정보산업연구소 연구전임교수

2013년 ~ 현재 광주과학기술원 전기전자컴퓨터공학부 연구조교수

<주관심분야: 정보검색, 데이터마이닝, 해양IT정보융합, 클라우드 컴퓨팅, IoT, 스토리지 시스템>

김종원



1997년 University of Southern California 연구 조교수

1999년 Technology Consultant for VProtect Systems Inc.

2000년 Technology Consultant for Southern California Division

of InterVideo Inc.

2001년 광주과학기술원 전기전자컴퓨터공학부 교수

2008년 ~ 현재 광주과학기술원 전기전자컴퓨터공학부 교수

<주관심분야: Future Internet, SDN & NFV, SDI>