

입출력 패턴에 기반한 스토리지 서버의 소비전력 절감 가능성 및 성능 평가 방법

박찬영 · 이재면 · 강경태 (한양대학교)

목 차	1. 서 론
	2. 데이터센터의 소비전력
	3. 데이터 접근 패턴 분석
	4. 서버의 성능 평가
	5. 결 론

1. 서 론

사용자가 생산하는 데이터의 양이 급격하게 증가하며 이에 대응하기 위한 서비스 제공자의 고민이 늘고 있다. 그림 1과 같이 Youtube 서비스 이용자들은 1분에 400시간의 동영상상을 올리며, 이를 위해 매일 1페타바이트(1,000,000GB)의 새로운 스토리지가 필요하다[1]. 사용자는 업로드된 자신의 자료가 언제까지나 보관되기를 원하지만, 이는 고스란히 서비스 제공자에게 큰 부담으로 작용하게 된다. 데이터 저장을 위해 데이터센터의 스토리지 서버를 끊임없이 증축해야 하며, 빠른 서비스를 위해 스토리지 서버를 활성 상태로 유지해야 하기 때문이다[2]. 이에 따라 데이터센터가 소비하는 전력은 지속해서 증가하게 되며, 데이터센터의 소비전력 절감이 매우 중요한 문제로 대두되고 있다.

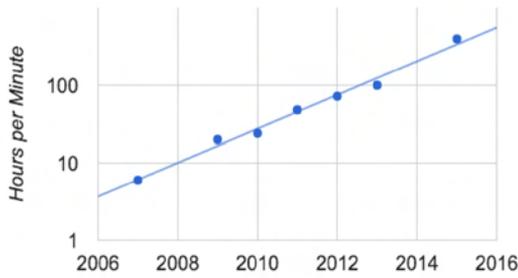
이 문제를 데이터의 접근 패턴 관점에서 해결하려는 방안이 있다. 이는 나이1)가 많은 데이터일수록 접근율이 낮다는 아이디어에 기초한다. 데이터의 나이와 접근율의 상관관계를 놓고 그래프를 그려보면, 나이가 적은 데이터가 서버 전체 접근율의 대부분을 차지하게 되는 tail latency 형태의 그래프가 그려진다[3]. 이때 접근 요청이 거의 없는 데이터를 콜드 데이터, 접근 요청이 빈번한 데이터를 핫 데이터라고 부르며, 콜드 데이터만을 모아 적절한 정책으로 관리하여 데이터센터의 소비전력을 절감할 수 있다.

이와 같이, 현재 산업계는 콜드 데이터에 집중하여 데이터센터의 소비전력을 절감하고자 한다 [4,5]. Open Compute Project²⁾ (OCP)에서는 콜드 데이터를 저장하기 위한 저전력 스토리지 서버(콜

1) 생성 시각으로부터 오래된 데이터인지 아닌지를 나이가 많다, 적다는 것으로 표현한다.

2) Open Compute Project는 하드웨어를 재구성하여 더욱 효율적이고, 유연하며, 확장성 있게 만드는 데 초점을 둔 컴퓨팅 인프라 구조 협력 커뮤니티이다[6].

1) 생성 시각으로부터 오래된 데이터인지 아닌지를 나이가 많다, 적다는 것으로 표현한다.



(그림 1) Youtube upload rate[1]

드 스토리지)의 명세서를 공개하였다[6]. Facebook은 콜드 스토리지를 디스크와 블루레이로 확장하고 있으며, 앞으로 플래시 메모리를 기반으로 한 콜드 스토리지를 내놓을 계획이라고 발표하였다[7]. 또한, Google은 2016년 USENIX가 개최하는 국제 학술대회에서 발표된 기고문을 통해 콜드 데이터를 위한 디스크 및 관리 도구의 필요성을 주장한 바 있다[1].

이처럼 스토리지 서버의 소비전력 절감을 위해 다양한 연구가 진행되고 있다. 하지만 소비전력을 절감하였다 하더라도 서비스가 요구하는 수준의 응답시간을 보장하지 못한다면, 연구 결과를 실제로 적용하는데 한계가 있다. 따라서 서버의 성능 평가 방법 역시 중요하다. 하지만 대규모 인프라에 초점을 맞춘 콜드 스토리지 연구는 학계에서 성능 평가를 진행하기에 규모 면에서 어려움이 있다. 이를 해결하기 위해 다양한 워크로드를 생성해낼 수 있는 벤치마킹 도구로 본 논문에서는 Yahoo! Cloud Serving Benchmark (YCSB)를 소개한다.

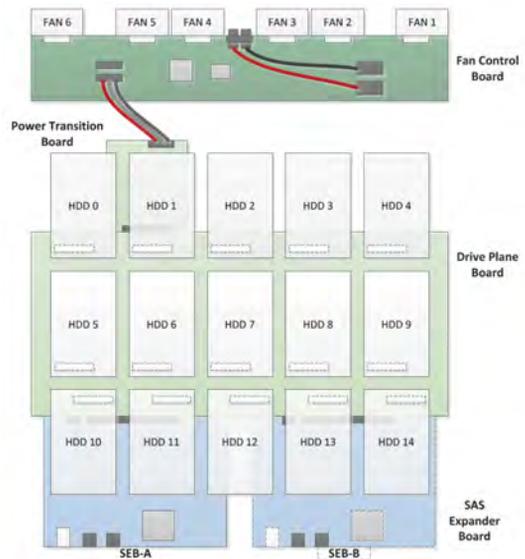
2. 데이터센터의 소비전력

데이터센터가 소모하는 전력은 어느 정도인가? 미국의 경우 이미 수년 전부터 데이터센터가 소비하는 전력이 뉴욕 시가 소모하는 전력의 2배를 넘어섰다[8]. 국내 데이터센터의 경우 2013년 기준 약

26억kWh의 전력을 소모하였다. 이는 한 달 동안 약 1,200만 가구가 사용하는 전력량과 비슷하다[9]. 또한, 새로운 데이터가 생성될수록 데이터센터가 소모하는 전력량은 늘어날 수밖에 없으며, 그 상승세 역시 가파른 실정이다.

이처럼 많은 양의 전력을 소모하는 데이터센터는 기업의 골칫거리이다. 우선 값비싼 전기세로 인해 유지비용이 비싸고, 친환경 에너지로만 전력을 공급받지 않는 이상 환경오염의 질책을 피할 수 없다[8-9]. 이를 해결하기 위해 다양한 산업계의 노력이 있었으며, 대표적으로 OCP의 콜드 스토리지가 있다. 콜드 스토리지는 에너지 절감을 위해, 연산 능력을 손해 보더라도 낮은 가격과 저전력으로 요청이 적은 콜드 데이터를 처리하는 스토리지 서버다.

OCP의 콜드 스토리지는 Open Vault³⁾의 기준(그림 2)을 따르며 소비 전력 최소화를 위해 몇 가지 요소를 변경하였다. 총 6개의 fan 모듈에서 2개를



(그림 2) Open Vault's major system components[6]

3) OCP에서 정의한 스토리지 서버를 모두 아울러 Open Vault라 한다.

제거하여 4개의 fan 모듈이 있다. 그리고 compute node의 사양을 낮추었으며 하나의 스토리지 서버에 하드디스크 2개만 동작하도록 하여 총 하드디스크의 개수와 상관없이 2개 분량의 전력을 소비하도록 설계되었다. 실제로 Facebook은 콜드 스토리지를 적용하여, 이전 대비 38% 이상의 에너지 효율과 24% 이상의 가격 경쟁력을 얻은 바 있다.

또한, 지속적인 플래시 가격의 하락으로 인하여 플래시를 이용한 콜드 스토리지 연구가 진행되고 있다. 플래시는 하드디스크보다 크기는 작고, 소비전력이 낮으며 진동과 발열이 적은 장점이 있다. 따라서 플래시로 콜드 스토리지를 구성하게 되면, 높은 집적도와 낮은 소비전력을 기대할 수 있다. 하지만 아직 하드디스크에 비해 4배 높은 가격으로 인해 가까운 시일 내 상용화는 어려울 것으로 예상된다.

데이터센터의 막대한 소비전력 문제는 관련 업계에 종사하는 사람뿐만 아니라 현대사회의 큰 주목을 받고 있으며[8-9], 이를 해결하기 위한 저전력 서버 연구가 산업계를 중심으로 활발하다. 특히, OCP를 중심으로 하드웨어 연구가 진행되고 있을 뿐만 아니라, 콜드 스토리지에서 최적의 성능을 내기 위한 소프트웨어 연구도 진행되고 있다. 소프트웨어 연구는 콜드 데이터의 분류로부터 출발하며, 이를 위해 데이터 접근 패턴의 분석이 선행되어야 한다[10,11].

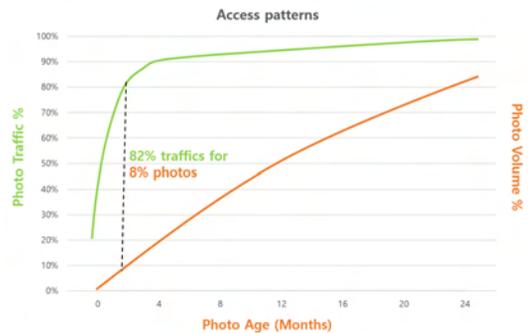
3. 데이터 접근 패턴 분석

데이터 접근 패턴 분석을 위해 데이터의 정렬이 필요하다. 이때 데이터의 나이를 정렬 기준으로 삼는 것이 일반적이다. 특히, 콜드 스토리지 운영 시 좋은 효과를 기대할 수 있는 소셜 네트워크, 동영상 공유 서비스와 같은 대규모 서비스의 경우 나이가 적은 데이터가 전체 트래픽의 대부분을 차지하는

형태의 패턴이 뚜렷하였다. 관련 사례로 Facebook과 LINE 서비스의 데이터 접근 패턴을 분석하였다.

3.1 Facebook의 데이터 접근 패턴

그림 3은 유명 소셜 네트워킹 서비스 제공사인 Facebook이 공개한 데이터 접근 패턴이다. 업로드된 사진 데이터를 나이로 정렬한 후 전체 트래픽에서 차지하는 비율을 나타낸 결과, 어린 나이의 데이터가 차지하는 트래픽 비율이 극단적으로 높았다. 이들이 공개한 수치에 의하면 2013년 기준 약 8% 사진 데이터가 전체 트래픽의 82%를 차지하였다.

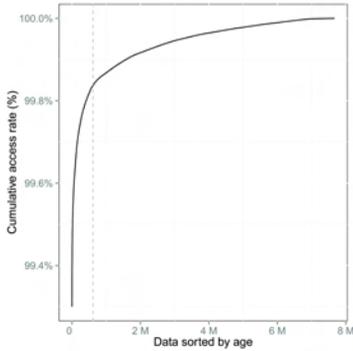


(그림 3) Facebook data access patterns, 2013[2]

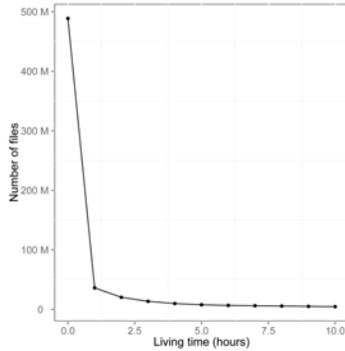
3.2 LINE의 데이터 접근 패턴

LINE은 무료 통화 및 메신저 애플리케이션이다. 사용자는 LINE을 이용하여 사진을 교환할 수 있다. 그림 3은 LINE에서 교환되는 사진 데이터를 나이 순으로 정렬하여 데이터 접근 패턴을 분석한 결과이다. 나이가 어린 8%의 데이터가 99.8%의 트래픽을 점유하였으며 앞선 Facebook과 유사한 형태를 보였다[12].

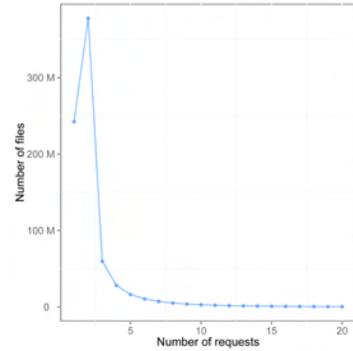
더욱 정확한 콜드 데이터 분류 정책을 위해 다양한 기준에서 데이터를 분석할 필요가 있다. 그림 5는 데이터의 생성 시각부터 마지막 요청 시각까지



(그림 4) LINE 데이터 접근 패턴[12]



(그림 5) LINE 이미지 파일 분포 - 활동시간 기준[12]



(그림 6) LINE 이미지 파일 분포 - 요청횟수 기준[12]

의 시차를 기준으로, 그림 6은 데이터의 요청 횟수를 기준으로 하여 데이터를 분석한 결과이다. 전자적 경우 데이터 생성 시각과 마지막 요청 시각의 차이(활동시간)가 4시간 이하인 데이터가 전체 데이터의 약 64%를 차지하였으며, 약 4%의 데이터가 활동시간 1시간에서 2시간 사이에 존재하였다. 이는 곧 활동시간 1시간 이후의 데이터는 접근 요청이 거의 없으며 콜드 데이터로 분류 가능함을 의미한다. 후자의 경우는 31.62%의 데이터가 한 번의 요청을 받고, 49.24%의 데이터가 두 번의 요청을 받는다. 따라서 약 80%의 데이터가 2번 이하의 요청 횟수를 가지며, 이는 일대일 대화방이 많은 LINE 서비스의 특징을 반영한 결과이다. 한번 접근된 사진 데이터는 장치에 저장되므로 추후 접근 시 재다운로드할 필요가 없다. 따라서 대화방의 인원수만큼 파일이 접근되면 콜드 데이터로 분류가 가능하다[12].

위 결과를 통해 Facebook과 LINE 모두 소수의 어린 데이터가 트래픽의 대부분을 차지하고 콜드 데이터가 많은 형태임을 확인하였다. 또한, LINE 서비스의 경우 더욱 정확한 콜드 데이터 선정을 위해 데이터의 활동시간, 요청횟수를 기준으로 파일의 분포를 분석하였으며, 다양한 분석 결과를 복합적으로 적용하여 콜드 데이터 분류가 가능했다. 분

류된 데이터를 콜드 스토리지에 저장 및 관리한다면, 데이터센터의 소비전력을 절감할 수 있다.

하지만 모든 서비스가 위와 같은 데이터 접근 패턴을 보이는 것은 아니다. Facebook, LINE과 같은 소셜 네트워크 서비스는 프론트 엔드부터 시간을 기준으로 한 콜드 데이터 분류에 유리하도록 설계되어 있다. 사용자에게 보이는 데이터는 기본적으로 시간이 지날수록 스크롤 아래로 밀려나게끔 설계되어있으며, 이러한 특성으로 인해 데이터 접근 패턴의 시간 의존성이 높아진다. 반면에 검색 엔진과 같이 시간 의존성이 높지 않은 서비스는 데이터 접근 패턴의 분석과 콜드 데이터 분류의 기준을 달리할 필요가 있다. 이처럼 다양한 서비스에 적합한 콜드 데이터 선정 기준은 콜드 스토리지의 적용 분야를 확장함에 있어 도전 과제로 남아있다.

4. 서버의 성능 평가

서버의 에너지 효율을 높이는 연구를 함에 있어서도, 서비스 운영을 위해 일정 수준의 응답시간 보장이 필요하다. 콜드 스토리지는 그 특성상 데이터 센터를 소유하고 대규모 서비스를 운영하는 곳에서 필요로 한다. 따라서 대규모 워크로드에서 서버의 성능을 쉽고 공정하게 비교할 방법이 필요하다.

〈표 1〉 YCSB 핵심 패키지의 워크로드[14]

Workload	Operations	Record selection	Application example
A-Update heavy	Read: 50% Update: 50%	Zipfian	Session store recording recent actions in a user session
B-Read heavy	Read: 95% Update: 5%	Zipfian	Photo tagging; add a tag is an update, but most operations are to read tags
C-Read only	Read: 100%	Zipfian	User profile cache, where profiles are constructed elsewhere (e.g., Hadoop)
D-Read latest	Read: 95% Insert: 5%	Latest	User status updates; people want to read the latest statuses
E-Short ranges	Scan: 95% Insert: 5%	Zipfian/Uniform	Threaded conversations, where each scan is for the post in a given thread (assumed to be clustered by thread id)

하지만, 데이터센터를 보유하기 힘든 학계는 콜드 스토리지 실험을 수행하는 데 한계가 있다. 또한, 실 서비스 워크로드가 없어 실험의 신뢰성을 보장하기 어렵다. 이에 벤치마킹 툴을 개발하여 워크로드 규모를 축소하거나 가상의 워크로드로 시뮬레이션하고 있다[13]. 이 접근 방법은 다양한 실험을 할 수 있는 장점이 있지만, 실제 서비스와의 유사성 문제를 안고 있다.

스토리지 서버를 실험하는 다양한 벤치마킹 툴이 존재한다[14-17]. 그 중 YCSB는 온라인 클라우드 서비스를 제공하는 데이터베이스 벤치마킹 툴이다[14]. 클라우드 서비스의 요청 패턴과 유사하도록 다중 클라이언트 기능을 제공하고, 워크로드를 사용자 요구에 맞게 수정할 수 있다. 특히, 요청 패턴을 5가지 정책(Constant, uniform, zipfian distribution, sequential, hotspot)에 따라 생성함으로써, 실제 서비스 요청과 유사한 패턴을 보일 수 있다.

오픈소스인 YCSB는 NoSQL 데이터베이스 전문 벤치마킹 툴이지만, Key-value Store 기반의 오브젝트 스토리지 기반 파일시스템도 연결 객체(binding)를 개발하여 실험 가능하다[18]. 현재 Ceph의 오브젝트 스토리지 RADOS binding을 공식 지원하고 있다[19].

사용자는 YCSB에 CRUD+S(읽기, 쓰기, 업데이트, 삭제, 스캔) 동작의 비중을 달리하여 다양한 요청 패턴을 실험할 수 있으며, 표 1은 YCSB에서 제공하는 대표 워크로드이다. 워크로드 별로 초당 명령 처리속도와 평균 대기시간 등 서버 성능 평가에 필요한 지표를 얻고 성능을 비교할 수 있다.

5. 결 론

데이터센터의 소비전력 절감이 TCO 관점의 문제를 넘어 매우 중요한 사회적 문제로 떠오르고 있다. 그린피스를 비롯한 환경단체로부터 데이터센터의 많은 전력 소비를 지적받고 있으며 세계 각국의 정부 부처는 데이터센터의 에너지 규제에 관심을 보이고 있다[20,21].

데이터센터의 소비전력 절감 가능성을 확인하기 위해 애플리케이션의 데이터 접근 패턴을 분석하였다. 그 결과 특정 서비스에서 나이가 어린 소수의 데이터가 트래픽의 대부분을 차지하였으며, 콜드 데이터에 집중한 소비전력 절감 연구가 진행되고 있음을 확인하였다. 또한, 더욱 정확한 콜드 데이터 선정을 위하여, 서비스별 특성을 고려한 데이터 접근 패턴 분석의 기준이 필요함을 확인하였다.

콜드 스토리지의 명세를 공개한 OCP를 중심으로 하드웨어 연구가 활발히 진행되고 있으며, 오픈 소스 분산파일시스템 커뮤니티를 중심으로 소비 전력 절감에 초점을 맞춘 연구 성과물들이 점차 발표되고 있다. 하지만 스토리지 서버 연구는 실제 서비스 규모가 방대하여 학계에서 연구하기 한계가 있다. 이를 위해 YCSB와 같은 벤치마킹 도구가 있으며, 스토리지 서버 연구의 성능 지표 검증이 가능함을 확인하였다. 그 결과, 학계에서도 관련 연구가 더욱 활성화될 것을 기대한다.

참 고 문 헌

- [1] Google, Disks for Data Centers, 2016.
- [2] OSCON – The OCP keynote by Jay Parikh, [Online]. Available: <http://www.oscon.com/oscon2013/public/schedule/detail/29558>
- [3] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon, "Analyzing the Video Popularity Characteristics of Large-Scale User Generated Content Systems," IEEE/ACM Transactions on Networking (TON), Vol. 17, no. 5, pp. 1357-1370, 2009.
- [4] E. Thereska, A. Donnelly, and D. Narayanan, "Sierra: Practical power-proportionality for data center storage," Proc. the 6th Conf. Computer Systems (Eurosys 11), pp. 169-182, 2011.
- [5] C. Albrecht, A. Merchant, M. Stokely, M. Waliji, F. Labelle, N. Coehlo, X. Shi, and E. Schrock, "Janus: Optimal flash provisioning for cloud storage workloads," Proc. USENIX Ann. Technical Conference (ATC 13), pp. 91-102, 2013.
- [6] Open Compute Project, [Online]. Available: <http://www.opencompute.org>
- [7] Open Compute Project US Summit 2016 by Jay Parikh, [Online]. Available: http://www.opencompute.org/wiki/Main_Page/Summit/2016US
- [8] NRDC, Data Center Efficiency Assessment, 2014.
- [9] 그린피스, "당신의 인터넷은 깨끗한가요?," 2015.
- [10] 이재면, 강경태, "OCP Cold Storage 테스트베드," 정보과학회 컴퓨팅의 실제 논문지, 제22권, 제3호, pp.151-156, 2016.
- [11] 이재면, 정해건, 원지웅, 강경태, "OCP Cold Storage 정책 평가를 위한 테스트 베드," 한국컴퓨터종합학술대회(KCC), 2015.
- [12] J. Lee, C. Song, and K. Kang, "Energy-efficient Storage Policy for Instant Messenger Services," in Proc. 5th IEEE International Conference on Big Data and Cloud Computing (BDCloud), pp. 38-44, 2015.
- [13] Y. Wang, M. Kapritsos, L. Schmidt, L. Alvisi, and M. Dahlin, "Exalt: Empowering Researchers to Evaluate Large-scale Storage Systems," Proc. the 11th USENIX Conference on Network Systems Design and Implementation (NSDI 14), pp. 129-141, 2014.
- [14] B. F. Cooper, A. Silberstein, E. Tam, R. Ramakrishnan, and R. Scars, "Benchmarking Cloud Serving Systems with YCSB," Proc. the 1st ACM Symposium on Cloud Computing, pp. 143-154, 2010.
- [15] M. Wittle, and BE. Keith, "LADDIS: The Next Generation in NFS File Server Benchmarking," USENIX Summer, pp. 111-128, 1993.
- [16] I. DRAGO, E. Bocchi, M. Mellia, H. Slatman, and A. Pras, "Benchmarking personal cloud storage," Proc. the 2013 conference on Internet measurement conference. ACM, pp. 205-212, 2013.

- [17] W. Sobel, S. Subramanyam, A. Sucharitakul, J. Nguyen, H. Wong, A. Klepchukov, S. Patil, A. Fox, and D. Patterson, "Cloudstone: Multi-platform, multi-language benchmark and measurement tools for web 2.0," Proc. CCA, Vol. 8, 2008.
- [18] J. Lee, C. Song, and K. Kang, "Benchmarking Large-scale Object Storage Servers," Proc. 40th IEEE Annual Conference on Computers, Software and Applications (COMPSAC), pp. 594-595, 2016.
- [19] YCSB - v0.10.0 Release note, [Online]. Available: <https://github.com/brianfrankcooper/YCSB/releases/tag/0.10.0>
- [20] NRDC, "America's Data Centers Are Wasting Huge Amounts of Energy," 2014.
- [21] 정상진, 김문구, "그린 데이터센터 표준화 및 인증제 동향," TTA Journal, Vol. 145, pp. 76-82, 2013.



이재면

이메일: jaemyoung@hanyang.ac.kr

- 2012년 한양대학교 컴퓨터공학과 (학사)
- 2012년~현재 한양대학교 컴퓨터공학과 석박사통합과정
- 관심분야: 지능형 모바일 컴퓨팅 플랫폼, 저전력 스토리지 시스템, 사이버피지컬 시스템



강경태

이메일: kt kang@hanyang.ac.kr

- 1999년 서울대학교 수학교산통계학과군 전산과학전공 (학사)
- 2001년 서울대학교 전기컴퓨터공학부 (석사)
- 2007년 서울대학교 전기컴퓨터공학부 (박사)
- 2008년 PostDoc. Res. Ass. Coordinated Science Lab., UIUC
- 2010년 PostDoc. Res. Ass. Dept. Computer Science, UIUC
- 2011년~현재 한양대학교 컴퓨터공학과 부교수
- 관심분야: 사이버피지컬 시스템, 모바일 컴퓨팅, 운영체제

저자약력



박찬영

이메일: chanyoung@hanyang.ac.kr

- 2016년 한양대학교 컴퓨터공학과 (학사)
- 2016년~현재 한양대학교 컴퓨터공학과 석박사통합과정
- 관심분야: 저전력 스토리지 시스템, 차량 시스템, 운영체제