

인메모리 시스템에서 백업을 위한 쓰기 간격에 따른 하드디스크의 성능 변화에 대한 연구

최찬호, 박세훈, 엄현상, 엄현영
서울대학교 컴퓨터공학과
e-mail : {chchoi, hseom, yeom}@dcslab.snu.ac.kr

Study on Throughput Variation of Hard Disk Drive according to Gap between Writes in In-memory System for Backup

Chanho Choi, Sehoon Park, Hyeonsang Eom, Heon Y. Yeom
Dept. of Computer Science Engineering, Seoul University

요 약

최근 서비스가 대규모화 되고 사용자 수와 시간 당 사용량이 폭발적으로 증가함에 따라 디스크 기반의 서버에서 메모리를 저장장치로 활용하는 인메모리 시스템이 개발되고 주목 받고 있다. 이러한 인메모리 시스템은 휘발성인 DRAM 을 사용하기 때문에 전력 공급이 차단되는 경우 크게 문제가 발생할 수 있는데, 이를 해결하기 위해서는 비휘발성 저장장치의 백업이 필수적이다. 본 논문에서는 비휘발성 저장장치 중 가장 저렴하고 일반적인 하드디스크를 활용하여 백업 시스템을 구축할 때, 디스크로 보내는 데이터를 최소화하면서도 디스크의 특성상 성능이 떨어지는 현상을 방지하기 위한 연구를 진행하였다.

1. 서론

과거부터 지금까지 컴퓨터의 점진적인 보급과 서비스의 발전으로 인해 유저가 폭발적으로 증가하고 서버에 들어오는 정보의 양도 막대하게 증가하고 있다. 대표적으로 Facebook, Amazon, Twitter 와 같이 대규모 서비스를 하는 기업에서는 데이터 센터를 구축하고 많은 양의 컴퓨터로 분산 시스템을 형성하여 서비스를 하고 있다[1][5]. 그러나 디스크 기반의 서버로는 많은 양의 데이터를 한번에 처리할 수가 없고, 반면 DRAM 의 용량이 증가함에 따라 단위 용량당 단가가 저렴해져서 마침내 인메모리 시스템(In-memory System)이 등장하기에 이르렀다. [6] 이렇게 등장한 인메모리 시스템은 디스크 기반 서버에 비해 수천 배 이상 빠른 성능을 자랑하지만 천재지변이나 다른 이유로 전력이 문제가 생기는 경우에 메모리가 휘발성 이기에 저장된 데이터를 모두 잃고 복구할 수 없게 되는 일이 발생한다. 물론 [6]와 같이 이를 대비하여 디스크로 백업을 하는 일을 행하고 있으나 그 속도가 메모리에 미치지 못해 여전히 위험은 남아있다.

하드 디스크 기반의 로깅 시스템[4]은 가격적으로 매우 효율적이므로 이를 이용한 시스템을 디자인 하는 것은 매우 중요한 부분이다.[2] 로깅 시스템에서는 메모리의 이미지를 디스크 상에 복사할 수 있도록 하는 부분이 핵심인데, 이 때 최상의 속도를 가지기 위해서는 디스크를 순차적 쓰기 작업(Sequential Write)으로 수행되어야 한다. 그러나 메모리 이미지를 항상

전체를 복사하는 작업은 시간도 오래 걸릴뿐더러 메모리에 주는 부담도 늘어나게 된다. 따라서 유효한 데이터만 복사하는 것이 중요한데 이는 무작위적 쓰기 작업(Random Write)를 유발하여 속도를 떨어뜨리게 되는 원인이 될 수 있다. 따라서 이 논문에서는 디스크에 이미지를 기록할 때 간격을 두고 기록하는 것이 어떤 결과를 가져올 수 있는지에 대해 실험하고 그 결과에 대해 기술하였다.

2. 디스크에 특성에 따른 쓰기 과정

디스크는 순차적 쓰기와 무작위적 쓰기에 심하게 다른 특성을 보여주는데, 이는 디스크에서 헤드의 기계적인 움직임이 필요하기 때문에 나타나는 특성이다. [3] 원하는 위치에 데이터를 쓰기 위해서는 헤드가 해당 위치로 움직이는 시간이 필요하기에 이로 인해 나타나는 차이는 순차적 쓰기에서는 150MB/sec 의 속도를 보이는 반면 무작위적 쓰기에서는 최악의 경우 1MB/sec 의 속도를 보일 정도로 이 효과는 심각하다.

이러한 특성 때문에 디스크에 메모리의 이미지를 복사하기 위해서는 순차적 쓰기로 기록해야 하는데, 이 경우 메모리에 가해지는 부담이 커지고 중복으로 복사되는 데이터도 많아지게 된다. 따라서 유효한 데이터만을 덧쓰게(Overwrite) 되는데 이 때 산재된 유효 데이터 사이에 많은 간격이 생겨나게 된다. 이 간격에 따라 성능이 어떻게 변화하는지 실험하였다.

3. 실험 환경

우리는 본 실험을 위해 파일 시스템이 만들어지지 않은 하드디스크에 파이썬을 이용해 특정 섹터에 바로 데이터를 기록할 수 있도록 하였다. 디스크는 Western Digital 사의 WD Caviar Black(1TB, 7200RPM, 64MB Buffer)를 사용하였고, 운영체제와 하드디스크의 버퍼링 효과를 피하기 위해서 실험 데이터는 총 10Gbytes로 실험하였다.

데이터를 쓰지 않는 빈 간격에 따라 성능이 어떻게 변화하는지 알아내기 위해서 먼저 4KBytes 블록을 쓰고 4KBytes를 뛰어넘는 일을 반복하여 총 10GBytes를 헤드가 움직이도록 하였다. 이후 실험에서는 4KBytes를 8KB, 16KB, 32KB 등으로 늘려가면서 실험을 진행하였다.

4. 실험결과

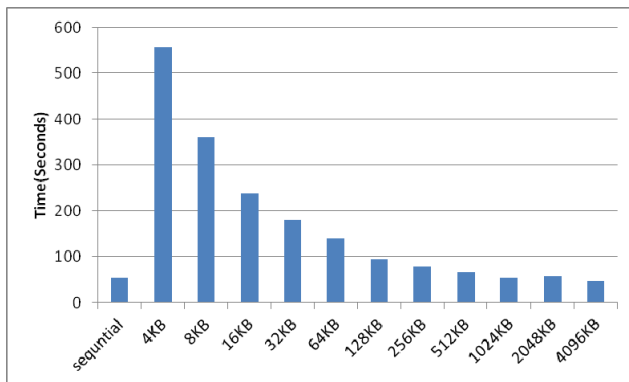


그림 1. 간격 별 10Gbytes 쓰기 시간의 차이

실험 결과는 그림 1 과 같이 나타난 것을 확인할 수 있다. 4KB 단위로 쓰고 건너뛰는 경우에는 순차적 쓰기에 비해 10 배 이상 느려지는 것을 볼 수 있다. 그러나 쓰고 건너뛰는 단위를 크게 할수록 이 차이는 점차 줄어들고, 약 512KB 이상부터는 순차적 쓰기와 크게 차이가 나타나지 않는 것을 볼 수 있다. 따라서 이는 인메모리 시스템에서 하드디스크로 백업을 할 시에 512KB 이하의 크기를 가지는 블록을 복사하는 것은 하드디스크의 성능을 떨어뜨릴 수 있다는 것을 의미한다.

5. 결론

이 논문에서 인메모리 시스템에서 하드디스크에 백업을 할 때, 어떤 간격을 두고 사용하여도 디스크를 최대한 활용할 수 있는지 실험을 통하여 확인하였다. 실험 결과 쓰기 간격이 512KB 이상이면 디스크 성능에 크게 변화를 미치지 않으므로 늘 512KB 이상의 쓰기와 간격을 두고 디스크에 보낸다면 보내는 데이터량을 최소화하면서도 디스크의 성능을 보장할 수 있으리라 생각된다.

차후에는 이러한 데이터를 기반으로 백업 시 디스크에 보내는 데이터의 패턴의 정책을 결정하여 디스크의 성능을 지속적으로 최상의 상태로 사용할 수 있도록 구성할 예정이다.

감사의 글

이 논문은 2011 년도 정부(교육과학기술부)의 재원으로 한국연구재단-차세대정보컴퓨팅기술개발사업의 지원을 받아 수행된 연구임. (No. 2011-0020521) 이 연구를 위해 연구장비를 지원하고 공간을 제공한 서울대학교 컴퓨터연구소에 감사 드립니다.

참고문헌

- [1] Kelsey, T. (2010). Social networking spaces from Facebook to Twitter and everything in between: A step-by-step introduction to social networks for beginners and everyone else. NY: Springer-Verlag
- [2] Ji-Yong Shin, Mahesh Balakrishnan, Lakshmi Ganesh, Tudor Marian, Hakim Weatherspoon. Gecko: A Contention-Oblivious Design for Cloud Storage, 4th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage'12), Boston, MA, June 13-14, 2012.
- [3] Chen BM, Lee TH, Venkataramanan V. Hard disk drive servo systems. London: Springer; 2002.
- [4] ROSENBLUM, M., AND OUSTERHOUT, J. K. The design and implementation of a log-structured file system. ACM Transactions on Computer Systems 10, 1 (1992), 26-52..
- [5] J. Dean and S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," in Proceedings of OSDI '04: 6th Symposium on Operating System Design and Implementation, San Francisco, CA, Dec. 2004.
- [6] Ousterhout, J., et al. The Case for RamClouds: Scalable High Performance Storage Entirely in DRAM. In SIGOPS OSR. December, 2009.