

하이브리드 저장 시스템을 위한 헤테로-레이드 기법에 대한 연구

변시우*, 허문행*

*안양대학교 디지털미디어학과

e-mail: swbyun@anyang.ac.kr, moonh@anyang.ac.kr

A Study on Hetero-RAID Scheme for Hybrid Storage Systems

Siwoo Byun*, Moonhang Hur*

*Dept of Digital Media, Anyang University

요 약

본 논문에서는 전통적인 RAID-1 미러링에 기반을 둔 헤테로-미러링이라는 하는 새로운 저장 장치 관리 기법을 제안한다. 헤테로-미러링 기반의 스토리지 관리는 SSD에서 발생 가능한 프리징 현상을 피하기 위한 쓰기-부하 밸런싱과 쓰기 지연 연산을 통하여 RAID-1 처리 성능을 개선한다.

1. 서론

지난 30년 전부터 하드 디스크(HDD: Hard Disk Drive)는 대부분의 데이터 서버에서 사용되는 절대적인 저장 미디어이다. 그러나 최근 중대형 데이터 서버에서의 저장 시스템으로서 최근 상용화된 Solid State Drive(SSD)가 주목 받기 시작하였다.

본 연구는 이러한 대용량 HDD 및 고속 SSD에 대하여, 저장 장치로서의 약점을 상호 보완하면서, 장점을 강화하여, 최종적으로 입출력 성능을 극대화할 수 있는 새로운 융합형 데이터 저장 시스템(Hybrid-RAID) 개발이다. Hybrid 저장 시스템 개발을 위한 최근의 기술 연구 및 기반 기술의 설계 및 구현에 대하여 논하며, 헤테로 미러링 기법을 제안한다.

2. 관련 연구

하드 디스크[6]에 대한 대안으로 최근에 주목받고 있는 저장장치가 바로 플래시 메모리[4,7]를 장착한 SSD이다. 결론적으로 용량과 가격측면에서는 하드 디스크가 매우 우수하고, 접근 성능 및 저전력 측면에서는 SSD가 매우 우수하다.[2]

이미 1990년대에 등장한 SSD는 하드 디스크에 비하여 많은 장점에도 불구하고, 그동안 가격 경쟁력이 매우 부족하여, 특수한 분야(군사, 항공, 우주, 선박등)에서만 주로 사용되었다. 그러나, 최근에는 플래시 메모리의 급속한 가격 하락에 힘입어, 서버, 휴대용 노트북등 고급 제품을 중심으로 기업 및 일반 소비자 시장으로 확대되고 있다.[5,8]

SSD의 문제점은 실제 사용자가 작업하다가, 저장 연산

이 집중되는 경우, SSD가 평소속도보다 상당히 느린 속도로 반응한다는 문제이다. 물론, 가끔씩 발생하지만, 하드 디스크에 비해서는 상당히 빈번히 발생하고, 홀딩 시간도 하드 디스크의 보다 훨씬 더 길다.

이 현상은 물론 읽기 연산에서는 거의 발생하지 않는다. 하지만, 소프트웨어를 설치하거나, 특히, 큰 콘텐츠를 복사 받는 경우 윈도우 운영체제가 심한 경우 수 초 이상씩 정지(흔히 프리징이라고 함)가 되는데, 이러한 체험 사례는 인터넷에서도 보고되어 있다.

3. 본 론

과거부터 하드 디스크 저장 시스템 업계에서는 저장 성능과 응답 성능의 문제해결을 위하여 여러 개의 디스크를 연동한 RAID[1,3] 기술을 사용하여 왔었다. 마찬가지로 SSD도 성능과 안정성을 모두 개선하여야 하므로, 이 RAID 기술의 활용이 필요하다.

이러한 하드 디스크 기반의 RAID 기술을 SSD에 그대로 적용해도, 어느 정도의 성능 개선과 응답 성능 향상을 얻을 수 있다. 그러나, 읽기 및 쓰기 성능이 대칭이 되는 하드디스크에 비하여, 쓰기가 매우 느려서 비대칭인 SSD를 하드디스크와 동일한 기법으로 저장 및 검색하면 심각한 I/O 문제가 발생한다. 즉, 플래시 메모리의 장점도 못 살리고, 단점도 극복하지 못하게 되어, 심각한 I/O 성능 저하로 나타날 수 있다. 단일 매체로 구성된 RAID는 그 구성 방법을 바꾸어도, 그 단일 매체의 자체의 본질적인 약점을 제대로 극복하기는 어렵다. 예를 들어, HDD만으로 새로운 RAID기법을 개발하

더라도, HDD의 구성이 달라질 뿐, 본질적인 기계적인 약점은 그 속에 그대로 내재되어 있기 때문이다. SSD의 경우도 마찬가지이다. 따라서, 서로의 약점을 보완해줄 성질이 다른 이질형(Heterogeneous)- RAID 기술이 필요하다. 즉, 한 매체의 약점이 다른 매체에서는 강점인 반대 특성을 활용하여, 그 약점을 흡수해야 한다.

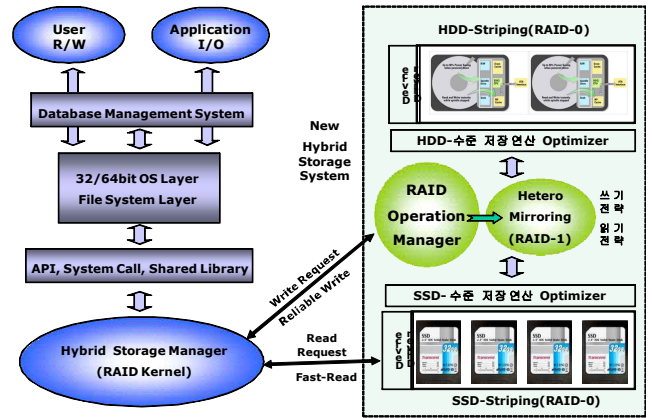
본 연구에서는 이러한 관점에서 기존의 RAID-1 기술을 개선하고, 고속 검색용 SSD의 약점을 보완하기 위하여, 대용량 HDD를 융합하여, 전체적인 시너지 효과를 낼 수 있는 새로운 헤테로 미러링(Hetero Mirroring) 기술을 제안 한다. 본 논문에서는 미러링된 스토리지 제어에서 읽기 연산은 어느 한쪽의 빠른 드라이브를 통하여 수행된다. SSD의 읽기 속도는 매우 빠르며 별 병목현상 문제가 없으므로 간략히 서술하며, 본 논문의 이슈인 저장 즉 쓰기 연산에 대하여 설명하고자 한다.

헤테로 미러링 기법을 활용하여 전술한 SSD 프리징 현상을 감소시킬 수 있다. 즉, 일반 부하시의 SSD의 응답시간이 피크치로 급상승할 때도, HDD의 응답시간은 정상적인 경우가 많다. 반대로 헤드 이동 등으로 HDD도 피크치로 응답시간이 길어질 수 있는데, 이때는 SSD이 응답시간은 정상적인 경우가 많다. 즉, 동일한 저장 연산을 수행하더라도 두 매체가 모두 비정상적으로 오랜 시간이 지연된다거나, 프리징에 걸리는 경우는 없었다.

이러한 상호 보완적인 특징을 잘 활용하여, 특성이 상반된 이질형 두 매체(SSD와 HDD)를 미러링하면, 어느 한 쪽 드라이브의 급격한 응답지연을 충분히 흡수할 수 있다. 기존의 하드 디스크 기반 RAID-1(미러링) 기술은 안정성과 처리 성능에 일대 혁신을 이룬 상용화된 매우 중요한 기술이다. 즉, 원래 기존 미러링의 목적이 동일한 매체를 동시에 복제하여 안정성을 높이는 것이라면, 헤테로 미러링은 좀 더 나아가서, 프리징의 충격까지 흡수시켜서 최종 성능까지 높이는 새로운 이질형 RAID 기술의 핵심이다.

또한, 전술한 데로, SSD를 빠른 입출력을 위하여 기존의 RAID에 그대로 적용하게 되면, 많은 성능과 안정성에 심각한 문제가 발생하게 된다. 그 주요 원인은 read/write 처리 시간이 비슷한 하드 디스크에 비하여, 플래시 메모리 SSD는 read/write 처리 시간의 매우 차이가 나는 극심한 비대칭적 구조를 가지고 있으므로, I/O의 밸런스가 제대로 맞지 않기 때문이다.

어떤 시스템도 전체적인 밸런스가 맞아야만 제대로 성능을 발휘할 수 있는 것과 같은 이치이다. 이러한 수배에 달하는 극심한 읽기/쓰기의 밸런스의 차이는 여러 개의 드라이브를 길게 일렬 연결(RAID-0, 스트라이핑)하여 어느 정도 극복할 수 있다. 즉, 일렬 배열 드라이브에 비례하여 성능은 점차 높아질 수 있다. 이러한 이질형 상호-보완 미러링과 더불어, 밸런스된 성능을 위하여 설계한 본 연구의 시스템 구성도는 아래와 같다.



[그림 1] Hetero-RAID 기반 Hybrid Storage

위 그림에서 사용자의 Read/Write 연산은 Hybrid Storage Manager를 통과한 후, 헤테로-미러링된 Reliable-Write 연산과 스트라이핑된 Fast-Read 연산으로 Redirection 된다. 또한, Hybrid Storage System의 RAID 기술과 더불어 프리징 장애를 방지하고, I/O성능을 더 가속시키려면, 프리징 징후를 빨리 파악하고 즉각 부하를 밸런싱할 수 있는 RAID-Operation Manager가 필요하다.

즉, SSD에서 소규모 데이터라도 많은 양으로 랜덤 쓰기를 아주 심하게 하면, 프리징 현상이 빈번히 발생되므로, 이러한 과밀 적체를 판단한 후 바로 적체를 부하 조절하여 피해갈 수 있도록 설계하여야 한다. 예를 들어 적체가 임계점을 초과하면서 응답시간이 급속히 느러지기 시작한다면, 오히려 HDD에 우선적으로 저장하고, SSD의 저장은 다른 대기 큐에 전환시키는 지연된 멀티큐 스케줄링이 효율적이다. 이렇게 지연된 SSD 큐잉은 추후에 쓰기 연산이 몰아져서 나오므로, 수명 연장의 효과도 볼 수 있다. 즉, 랜덤 쓰기가 모여져서 순차쓰기로 유도되는데, 플래시 메모리 SSD도 순차쓰기가 빠르며, 특히 작은 크기의 파일일 경우 더 효과적이다. 이는 특히 중대형 데이터베이스 서버나 Multitasking 작업이 많은 사용자 환경에서도 예측가능하고 고르게 안정된 응답시간을 더 높은 수준으로 보장해 줄 수 있다.

제안 기법인 Hetero-Mirroring의 수행 알고리즘은 다음과 같다.

- ① 사용자의 프로그램이 헤테로-미러링된 하이브리드 스토리지에 대한 읽기, 쓰기 연산(operation)을 요청한다.
- ② 사용자가 입출력 요청을 접수되면, 요청된 연산은 RAID-1 연산으로 전환된다. 이를 관리하는 Operation Manager는 입출력 요청연산을 자신의 연산-큐의 마지막에 삽입한다. 추후에 지금 요청한 연산이 완료되면 결과 값을 최종적으로 사용자 프로그램에 리턴 한다.
- ③ 연산-큐에서 한 연산씩 차례로 꺼낸 후 연산의 종류를 분류한다. 읽기 연산이면, 4)HM_READ을 수행하며, 쓰기 연산이면 5)HM_WRITE를 수행한다.
- ④ HM_READ(Hetero-Mirroring_READ):
1) 먼저, 읽기 속도가 빠른 SSD에 읽기연산을 포워드한

다. 기존의 미러링의 경우 두 디스크에 반반씩 나누어서 읽기가 가능하지만, SSD가 HDD보다 3배 이상 읽기 속도가 빠르므로, SSD가 전체를 읽는 것이 유리하며, 상대적으로 전력절감 효과도 크다.

- 2) SSD로부터 읽기 작업이 성공하면, 판독한 값을 Operation Manager에게 리턴 한다.
 - 3) 만일, SSD로부터 읽기 오류가 발생하면, HDD에 동일하게 읽기 요청을 한 후, 판독한 값을 Operation Manager에게 리턴 한다. 발생한 오류를 시스템에 보고 한다.
- ⑤ HM_WRITE(Hetero-Mirroring_WRITE):
- 1) 미러링은 SSD와 HDD 모두에 동일한 데이터를 저장해야하므로, 두 드라이브 모두에게 저장 명령을 수행한 후 기다린다.
 - 2) 두 드라이브 중에 한 개라도 저장 연산의 수행이 정상적으로 완료되었다는 이벤트가 들어오면, 처리 시간(wt)을 기록하고, Operation Manager에게 저장 완료 메시지를 리턴 한다.
 - 3) 나머지 한 드라이브의 완료 이벤트를 설정된 시간(wait_limit: wt*1.3로 기본 값으로 함)만큼 더 기다린다. wait_limit은 먼저 완료한 저장 속도에 대한 상대적인 시간 값으로, 클수록 동기화 간격은 짧아지나, 응답 속도는 더욱더 느려진다. 이를 설정하지 않으면, 홀딩현상이 발생할 경우 저장 연산에 대한 응답 대기 시간이 심각하게 길어진다.
 - 4) 만일, 해당 드라이브가 wait_limit을 초과하면, 이 드라이브에는 일시적으로는 더 이상 저장 연산을 큐잉하지 않고, 대신 지연 저장을 위한 Delay_Queue에 쌓이게 된다.
 - 5) 만일, 해당 드라이브로부터 쓰기 오류가 발생하면, 이 오류를 시스템에 보고 한다. 심한 경우 드라이브 장에 오류이면, 장애 디스크의 복구 요청을 RAID에 전달하고, 현재의 미러링 모드를 단일 드라이브 모드로 전환하여 계속 운영을 가능하게 한다.
- ⑥ HM_DELAY_WRITE(Hetero-Mirroring_DELAY_WRITE):
- 1) 해당 드라이브가 비지 상태이면 계속 대기하고, 아니면 Delay_Queue에 큐의 맨 앞에 있는 연산을 선택하여, 해당 드라이브에서 지연된 쓰기 연산을 가져 온다.
 - 2) 해당 드라이브에 지연된 쓰기 연산을 수행시킨다.
 - 3) 쓰기 연산 작업이 성공하면, 완료 메시지를 Operation Manager에 리턴 한다.
 - 4) 큐에서 지연된 연산이 있으면, 큐의 내용이 없을 때까지 1)번 이하를 반복한다. 결과적으로, 지연 연산 큐에 쓰기 연산들이 모이므로, 유희시 한꺼번에 연속해서 저장할 수 있다. 따라서, SSD의 경우 모아쓰기를 유도하는 부수적인 효과가 발생하여, SSD이 저장효율과 전력절감 효과가 발생한다. HDD의 경우에도 랜덤한 임의 쓰기 패턴이 모여서 한꺼번에 순차쓰기 효과가 있다.

4. 결 론

기존의 노트북, UMPC, 데스크탑, 서버에 탑재된 하드디스크의 새로운 대체품으로 SSD가 급부상하고 있다. 본 연구에서는 기존의 하드 디스크 기반 저장 시스템에서 성능과 안전성을 높이기 위하여 사용하고 있는 핵심 RAID 기술을, 주요 RAID 레벨들에 대하여 비교 분석하였고, 새로운 SSD의 매체 특성에 맞게 보완하여 개선된 설계 방안을 제시하였다. 또한, 전통적인 RAID-1 미러링에 기반하여 SSD의 특성을 고려한 헤테로-미러링이라는 하는 새로운 저장 장치 관리 기법을 제안하였다.

제안한 헤테로-미러링 기법은 SSD에서 발생하는 느린 쓰기 속도와 프리징 현상을 회피하기 위하여, 저장 매체간의 심각한 저장 속도의 차이를 고려하여, 일시적으로 빠른 저장 매체를 우선 저장하고, 느린 저장 매체는 지연 저장을 수행하게 하여 전체적인 성능을 개선하게 된다.

참고문헌

- [1] 김호진,황인철,맹승렬, 윤현수, "SDIO에서 RAID 레벨 5의 효율적인 구현," 정보과학회 학술발표대회, 제31권 제1호, pp. 64-66, 2004
- [2] 배영현, "고성능 플래시 메모리 SSD 설계 기술", 정보과학회지 vol 25, No. 6, pp. 18-28, 2007.6
- [3] 최귀열, 박계원, "지능형 I/O 구조를 갖는 RAID 시스템의 성능향상을 위한 연구," 정보통신학회지 제10권 11호, pp. 1989-1995, 2007
- [4] Chanik Park, Jaeyu Seo,Dongyoung Seo, Shinhan Kim, Bumsoo Kim, "Cost-Efficient Memory Architecture Design of NAND Flash Memory Embedded Systems", 21st International Conference on Computer Design, San Jose, California, pp. 474-479, October 13-15, 2003
- [5] Li-Pin Chang, Tei-Wei Kuo, "An Efficient Management Schemefor Large-Scale Flash Memory Storage Systems," Proc. of ACM SAC'04, pp. 862-868, Nicosia, Cyprus, March 14-17, 2004
- [6] SpinPoint T Series, <http://www.samsung.com/Products/HardDisk/SpinPointTSeries/index.asp>, 2007.
- [7] What is Flash?, <http://www.samsung.com/Products/Semiconductor/Flash/WhatisFlash/Flash.htm>, 2007.
- [8] Yim K., "A Novel Memory Hierarchy for Flash Memory Based Storage Systems", Journal of Semiconductor Technology and Science, 5:(4), pp. 262-269. Dec. 2005