고속 RAID 저장 시스템를 위한 레벨1-미러링 기법의 개선

변시우*, 허문행*
*안양대학교 디지털미디어학과
e-mail: swbyun@anyang.ac.kr, moonh@anyang.ac.kr

An Improvement of Level-1 Mirrorring Scheme for High-Speed RAID Storage Systems

Siwoo Byun*, Moonhang Hur*
*Dept of Digital Media, Anyang University

요 약

고속 저장 장치로 사용되는 플래시 메모리 기반의 SSD와 전통적인 HDD에 대한 기초 연구와 개선점에 대하여 논하였다. 고속 SSD는 낮은 전력소모, 빠른 데이터 읽기속도 등의 특징으로 데스크톱 및 서버용 데이터베이스의 핵심 저장 요소가 되었다. 하지만, 빠른 읽기 연산에 비하여 상대적으로 느리거나 프리징이 있는 SSD의 쓰기 연산 특성을 고려하여 HDD와 RAID에 기반을 둔 기존의 전통적인 스토리지 관리 기법을 개선할 필요가 있다. 이를 위하여, 본 논문은 전통적인 RAID-1 미러링에 기반을 둔 복합적 미러링 기법을 제안한다. 제안 기법은 SSD에서 발생 가능한 프리징 현상을 피하기 위한 쓰기-부하 밸런싱과 쓰기 지연 연산을 통하여 RAID-1 처리 성능을 개선한다.

1. SSD와 HDD

이미 1990년대에 등장한 SSD(Solid State Drive)는 하드 디스크에 비하여 많은 장점에도 불구하고, 그동안 가격 경쟁력이 매우 부족하여, 특수한 분야(군사, 항공, 우주, 선박 등)에서만 주로 사용되었다. 그러나 최근에는 플래시 메모리의 급속한 가격 하락에 힘입어, 서버나 휴대용 노트북등 고급 제품을 중심으로 기업 및 일반 소비자 시장으로 확대되고 있다. 하지만 아직도 하드 디스크(HDD: Hard Disk Drive)는 대부분의 데이터 서버에서 사용되는 절대적인 저장 미디어이다. 본 연구에서는 이러한 고속 SSD에 대하여, 기존의 저장 장치로서의 약점은 상호 보완하여, 입출력 성능을 극대화할 수 있는 융합형 저장 기술인 헤테로-미러링(Hetero -Mirroring) 기법을 제안한다. 본 기법은 SSD의 신뢰성 측면에서 가장 중요한 레벨인 RAID-1에 대한 응답시간 및 처리 성능을 개선한다.

2. 관련 연구

최근들어 하드 디스크[6]에 대한 대안으로 최근에 주목받고 있는 저장장치가 바로 플래시 메모리[4,7]를 장착한 SSD이다. 결론적으로 용량과 가격측면에서는 하드 디스크가 매우 우세하고, 접근 성능 및 저전력 측면에서는 SSD가 매우 우세하다.[2,5] SSD의 문제점은 실제 사용자가 작업하다가, 저장 연산이 집중되는 경우, SSD가 평소속도보다 상당히 느린 속도로 반응한다는 문제이다. 물론, 가끔씩 발생하지만, 하드 디스크에 비해서는 상당히 빈번히 발생하고, 홀딩시간도 하드디스크 보다 훨씬 더 길다.

3. 복합 미러링 기법

여러 개의 하드 디스크를 연동한 RAID[1,3] 기술은 성능 개선의 장점으로 꾸준히 사용되었다. 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 성능

이 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 개발하더라도, HDD의 구성이 달라질 뿐, 본질적인 기계적인 약점은 그 속에 그대로 내재되어 있기 때문이다. SSD의 경우도 마찬가지이다. 따라서, 서로의 약점을 보완해줄 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.

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 사용된다.

4. 성능 평가

본 연구에서 제안된 헤테로 미러링 기법의 저장 성능을 검증하기 위하여 실제 환경에서 작업 부하를 주고, 그 실험 결과를 분석해 보았다. 본 실험 수행 시 제안한 헤테로 미러링기법(Hete-Mr)과 비교된 저장 관리 기법은 SSD를 1개만 사용하는 단일디스크 저장 기법(SSD-1)과, 가장 안정적인 SSD RAID 구성인 미러링 기법(SSD-Mr)이다.

4.1 실험 환경

실험이 수행된 하드웨어 환경은 최대한 일반 사용자 환경에 가깝게 구성하였다. DELL 펜티엄4 CPU와 메인 메모리 2G, 운영체제는 현재 가장 많이 사용하고 있는 윈도우 XP를 사용하였다. 언어는 비주얼 C++를 사용하였으며, 부수적으로 작업 부하 생성 모듈과 물리적인 실험후 통계 자료 수집의 편의상 CSIM 시뮬레이션 라이브러리를 사용하였다. 성능 측정 부분은 모두 SATA로 연결된 SSD와 HDD를 통하여 검출되었다.

본 실험에 앞서서, 실험용 샘플이 되는 데이터가 필요한데, 실제 사용자의 윈도우 환경에서 사용 중인 모든 파일을 검색한 후, 사용 빈도가 가장 높은 순으로 나열해 보면, 96%이상이 500KB 미만의 작은 크기의 스몰 파일들이다. 이를 근거로 하여, 10KB부터 500KB 사이의 10KB 단위로의 랜덤 파일을 생성하여 미리 저장하여 두었다.

실험을 위한 저장 워크로드 생성 및 결과 분석 모델이 필요한데, 이는 CSIM에서 제공되는 폐쇄형 큐잉 모델(Closed Queuing Model)과 통계 라이브러리를 사용하였다. 이 모듈들을 통하여 초당 일정한 수의 쓰기 오퍼레이션을 생성하고, 이와 연관된 처리 시간과 성능을 측정하면 된다. 본 실험의 주요 성능 평가 지표는 표1과 같다.

[표 1] 성능 평가 지표

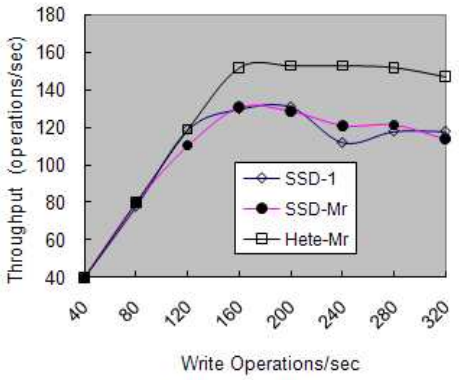
지 표	설 명	단위
operation throughput	초당 수행 완료된 저장 연산의 수	operations sec.
response time	저장 연산이 발생한 후 수행완료까지의 소요 시간	millisec

본 실험에서 생성되는 저장 오퍼레이션의 수는 초당 40개에서부터 320개까지 40개 단위로 8단계로 변화시켜 보았으며, 이는 시스템에 가해지는 작업 부하를 의미한다.

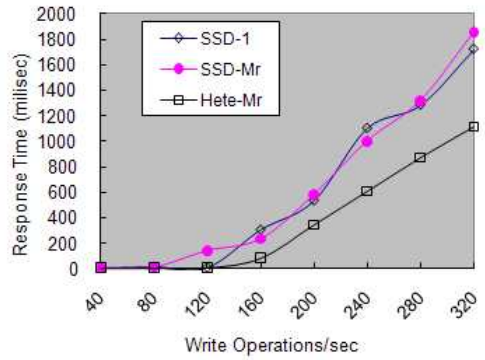
4.2 실험 결과 및 분석

본 실험은 기본적으로 하드 디스크나 SSD를 탑재한 일반적인 사용자 시스템에서 부과한 저장 워크로드, 즉 초당 발생된 저장 연산의 수가 시스템 성능에 어떤 영향을 미치는 지를 분석하기 위함이다. 그림1은 초당 발생된 저장 연산의 수의 증가에 따른 저장 연산의 처리치를 표시한 그래프이다. 발생된 초당 저장 연산의 수가 늘어날수록 점차로 연산의 처리 결과치가 증가함을 알 수 있다. 또한, 전반적인 저장 연산의 처리 성능을 측정한 결과, 제안기법인

Hete-Mr가 SSD-1과 SSD-Mr보다 높게 나타났다.



[그림 1] 초당 저장 연산 처리치의 비교



[그림 2] 저장 연산의 응답 시간의 비교

- 그림1과 그림2의 가로축은 저장 시스템에 가해지는 저장 연산의 부하를 의미하며, 그림 1의 세로축은 그 부하에서의 처리된 저장 연산수를 의미하며, 그림 2의 세로축은 저장연산에 걸린 응답시간을 의미한다.
- 그림1에서 보면, 초당 저장 연산의 수가 대략 160개를 넘으면서 각 저장 기법의 성능이 점점 낮아진다. 이는 초당 저장 연산수의 증가에 의한 플래시 메모리 기반 SSD 스토리지에서의 I/O 연산 집중화가 성능에 영향을 크게 미치는 주요 요소임을 의미하며, 이 수치 이상으로 워크로드를 활성화시키는 것이 성능 향상에 도움이 되지 않음을 의미한다. 특히 SSD를 구성하는 플래시 메모리의 쓰기 연산의 느린 속도와 수반 되는 소거 연산이 주요 요인이다. 그래프에서 보면 저장 연산의 부하가 적은 시작 구간에서는 SSD의 느린 쓰기와 소거 연산이 시스템 안에서 수용되어 성능 저하를 일으키지 않는다. 하지만, 시작 구간을 지나면서 생성되는 저장 연산의 수가 증가하므로, 플래시 메모리에 하위 레벨의 쓰기 연산의 수도 증가하게 된다. 따라서 중간 구간 이후에서는 저장 연산들이 SSD 컨트롤러의 작업 큐에 과도하게 쌓이면서, 저장 연산 처리가 지체되며 결과적으로 성능이 서서히 저하되게 된다.
- 하지만, 동일한 워크로드의 조건에서도 제안한 Hete-Mr 저장 기법이 SSD-1나 SSD-Mr 기법에 비하여 성능이 상대적으로 더 높다. 전체 구간에서 평균하면, Hete-Mr가 기존 기법보다 17.8% 정도 처리 성능이 높다. 특히, 최대 저장 연산 처리치는 Hete-Mr가 147개로 SSD-Mr의 114개 보다 28.9% 정도 더 높다. 그 이유는 위의 저장 성능 저하의 주요 원인인 SSD의 쓰기 연산의 부담에 의한 일시적인 홀딩현상을 헤테로 미러링기법은 일정 시간 비율을 초과하여 심각하게 길어지는 매체의 경우에는 일단 한 매체에 저장후 리턴하고, 늦어지는 매체는 지연 저장을 수행함으로써 전체적인 대기시간을 감소시킬 수 있었기 때문이다. 또한 SSD의 플래시 메모리 영역에 대하여, 쓰기 연산이 뒤로 지연되어 모아쓰기를 유도하는 효과도 발생하므로, 플래시 메모리의 수명이 조금 더 연장되며, 소거 연산을 위한 전력소모도 좀 더 줄이는 부수적인 효과도 기대할 수 있다.
- 그림2에서 사용자 프로그램이 요청한 저장 연산에 대한 각 스토리지 관리 기법의 응답시간을 비교해보면, Hete-Mr, SSD-1, SSD-Mr 순으로 우수하게 나타났다. 단일 SSD를 사용하는 SSD-1 기법에 비하여, SSD를 미러링한 SSD-Mr가 응답시간이 5% 더 늦게 나타났는데, 이는 SSD-Mr는 두 개의 SSD에 모두 즉, 2번의 저장이 완료되어야 하기 때문이다. Hete-Mr는 SSD-1에 비하여, 38% 정도 응답성능이 빠른데, 이는 전술한 데로, Hete-Mr는 HDD나 SSD 중에 어느 한쪽이 먼저 저장되면 바로 저장 연산을 완료시키므로 SSD의 심각한 홀딩현상을 회피가능하기 때문이다. 이러한 홀딩현상이 풀리면, 지연되었던 저장연산이 최종적으로 완료되게 된다.

5. 결 론

본 연구에서는 고속 SSD의 매체 특성에 맞게 보완하여 개선된 RAID-1 설계 방안을 제시하였다. 또한, 전통적인 RAID-1 에 기반하여 SSD의 특성을 고려한 헤테로-미러링이라는 하는 새로운 저장 장치 관리 기법을 제안하였다. 제안한 헤테로-미러링 기법은 SSD에서 발생되는 느린 쓰기 속도와 프리징 현상을 회피하기 위하여, 저장 매체간의 심각한 저장 속도의 차이를 고려하여, 일시적으로 빠른 저장 매체를 우선 저장하고, 느린 저장 매체는 지연 저장

을 수행하게 하여 전체적인 성능을 개선하게 된다.

참고문헌

[1] 김호진,황인철,맹승렬, 윤헌수, "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 frash? 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