

# SSD를 위한 파일시스템의 벤치마크

이성진<sup>o</sup>, 원유집

한양대학교 전자컴퓨터통신공학과

[insight@hanyang.ac.kr](mailto:insight@hanyang.ac.kr), [yjwon@hanyang.ac.kr](mailto:yjwon@hanyang.ac.kr)

## Performance Benchmark of Filesystem for Solid State Drive

Seongjin Lee<sup>o</sup>, Youjip Won

Department of Electronics and Computer Engineering, Hanyang University

### 요 약

파일시스템과 I/O 워크로드에 따라 Solid State Drive (SSD)의 성능은 변하는 것이 일반적인 사실이지만 파일시스템과 SSD와의 상관 관계는 명확히 알려지지 않았다. 파일시스템과 SSD 간의 상관관계 그리고 성능을 이해하기 위해 SSD와 4개의 파일 시스템을 IOzone 벤치마크에서 다양한 I/O 크기와 direct I/O 모드에서 실험을 하였다. Nilfs2를 제외한 모든 파일 시스템의 성능이 일정하지 않은 것으로 보이고 2MB 이상의 파일 크기에서는 100MB/s 정도의 성능을 보이고 있다. 반면 Nilfs2의 경우는 레코드 크기, I/O 단위가 작을 경우 성능 역시 낮고 I/O의 크기가 커질 경우 그 성능이 같이 증가하고 있는 것을 볼 수가 있다.

### 1. 서 론

플래시 메모리가 스토리지 서브시스템으로 사용됨에 따라 컴퓨터 시스템에 병목의 원인이었던 저장장치 기술이 빠르게 발전하였다. 플래시 메모리는 비휘발성 반도체로서 하드디스크에 비해 빠른 접근 속도와 저전력, 저소음 등 많은 장점을 갖고 있는 소자로서 차세대 저장 장치로서 각광받고 있다. 많은 장점에도 불구하고 아직까지 대중화가 되지 못한 몇 가지 이유는 아직 GB/\$가 높다는 것과 한정된 지우기 횟수 그리고 덮어쓰기가 불가한 특성 때문이다 [1]. 낸드 플래시의 각 셀은 쓰기 횟수가 제한적이다. 셀 당 저장 비트가 많아지면 SLC에서 십만 번 쓸 수 있었던 쓰기 횟수는 MLC에서 만 번으로 줄어든다. 쓰기 횟수 이상으로 저장을 하게 되면 신뢰성은 저하된다 [2].

대부분의 파일시스템은 단일 HDD의 성능을 높이기 위해서 개발되었기 때문에 HDD를 기반으로 한 스토리지 시스템에서는 성능 개선을 위한 최적화 기술들이 스토리지 시스템의 최적의 성능을 끌어올리지 못한다고 Stein [3]은 주장한다. 이와 같은 현상은 복잡한 FTL을 갖고 있는 SSD에도 동일하게 적용이 된다. 현재 알려진 대부분의 파일시스템은 SSD의 문제점을 고려하지 않고 개발이 되었다. SSD는 읽기/쓰기(페이지 단위) 그리고 지우기(블록 단위)의 단위가 다르기 때문에 log-structured 파일시스템이 아닌 경우 write amplification 효과로 성능이 저하될 수가 있다.

### 2. 실험 환경

파일시스템과 SSD 간의 성능 측정 및 비교를 하기

위해서 Intel Quad Core 2.4Ghz 4MB 캐쉬와 4GB DRAM 그리고 리눅스 커널 2.6.32-24을 사용하여 실험하였고 프라이어리 디스크는 WD1600 HDD를 사용하였다.

평가에 사용된 파일 시스템은 Ext3, Fat32, Nilfs2, 그리고 XFS로 총 4개이지만 본 논문에서는 지면 관계상 Ext3와 XFS의 성능을 나타낸다. Ext3d와 Fat32는 일반적으로 잘 알려진 파일시스템이고 Nilfs2는 New Implementation of Log-structured File system으로 리눅스에서 동작하도록 포팅된 로그 기반의 파일시스템이다. XFS는 확장성이 뛰어난 것으로 알려진 저널링 파일시스템이다.

사용된 벤치마크 툴은 IOzone으로 파일시스템의 순차 읽기/쓰기 임의 읽기/쓰기 성능을 평가하는데 사용된다. IOzone의 장점 중의 하나는 사용자가 I/O의 크기와 파일의 크기를 정의 할 수 있다는 점이다. 사용된 SSD 제품은 총 다섯 개로 표 1에 해당 제품의 간단한 성능명세와 함께 나타내었다.

SSD의 성능을 측정하기 위해서 I/O 크기와 파일의 크기를 변경하여 실험을 하였다. IOzone 벤치마크를 사용하여 SSD의 순차와 임의 읽기 그리고 쓰기 성능을 평가 한다. I/O의 크기는 4KB에서 1MB까지 각 파일 크기에 대해서 실험하였다. 파일 크기는 64KB에서부터 256MB까지 가장 작은 파일 크기에서 두 배씩 증가하도록 하였다.

O\_SYNC와 O\_DIRECT 함수를 사용하면 파일을 Synchronous I/O와 Direct I/O 모드로 Open하게 된다. Direct I/O를 사용하게 되면 파일시스템에서 스토리지를 접근할 때 성능향상을 위한 버퍼공간을 거치지 않고 바로 접근한다. 스토리지는 성능 향상을 위해 read-ahead를 매우 적극적으로 하기 때문에

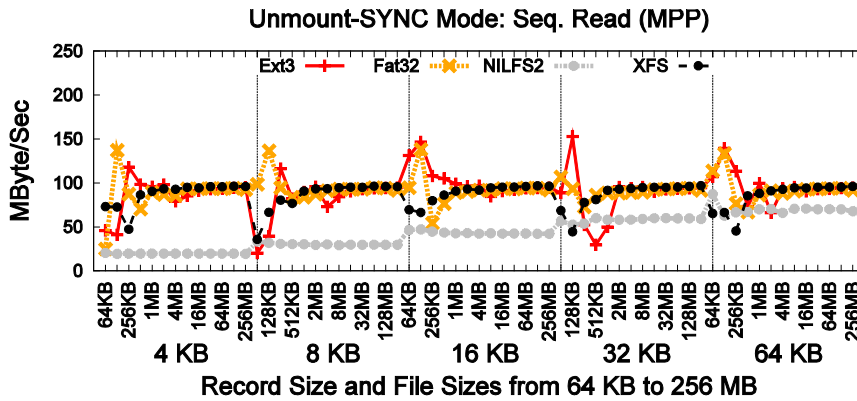


그림 1 순차 읽기 (Direct IO 모드): 파일 크기는 64KB에서 256MB까지 두배씩 증가하고 레코드의 크기는 4KB로 고정

표 1 SSD 성능명세

| SSD      | 세대 | 인터페이스 | 용량    | 타입  | 읽기      | 쓰기      | 채널   |
|----------|----|-------|-------|-----|---------|---------|------|
| 삼성 MPP   | 1  | SATA2 | 64GB  | MLC | 80MB/s  | 80MB/s  | 4ch  |
| 인텔 X25-M | 2  | SATA2 | 80GB  | MLC | 250MB/s | 70MB/s  | 10ch |
| 삼성 MXP   | 2  | SATA2 | 128GB | MLC | 220MB/s | 185MB/s | 8ch  |

스토리지 본연의 성능을 평가하기가 어렵다. Direct I/O의 경우 read-ahead를 사용하지 못하도록 설정하였다. O\_SYNC를 사용하는 Buffered I/O는 스토리지에 쓰기 전에 Buffer cache를 거치게 된다. 하지만 매 실험이 서로 독립적으로 평가 될 수 있도록 리마운트를 하여 버퍼 캐시를 해제하는 과정을 포함하였다.

### 3. 실험 결과

그림 1은 삼성 MPP SSD에서 실험 한 것으로 본 연구의 동기 역할을 하는 도표이다. 표 1에는 사용된 SSD의 성능 명세를 나타낸다. 실험 결과는 IOzone 벤치마크[4]를 사용하여 얻은 결과로서 레코드 크기를 4KB에서 64 KB 까지 증가를 하고 파일 크기를 64KB에서 256MB 까지 증가를 하였고 SYNC 모드를 사용하였다. SYNC 모드라는 것은 O\_SYNC() 명령을 사용하여 I/O를 처리하는 것으로 버퍼를 사용하는 모드를 말한다. IOzone의 결과는 레코드 크기에 따라 그룹 지어 표현하였다. 이 실험에서 사용된 파일 시스템은 Ext3, Fat32, Nilfs2 그리고 XFS 이다.

IOzone으로 측정된 읽기 성능을 보면 1MB 이하의 작은 파일 크기에서는 Nilfs2를 제외한 모든 파일 시스템의 성능이 일정하지 않은 것으로 보이고 2MB 이상의 파일 크기에서는 100MB/s 정도의 성능을 보이고 있다. 반면 Nilfs2의 경우는 레코드 크기, I/O 단위가 작을 경우 성능 역시 낮고 I/O의 크기가 커질 경우 그 성능이 같이 증가하고 있는 것을 볼 수가 있다. 이와 같이 파일 시스템의 종류에 따라 그리고 SSD의 종류에 따라 성능의 차이가 크게 나타나는 것을 볼 수가 있다.

### 4. 결론

본 논문에서는 파일 시스템과 I/O 워크로드와 SSD의 성능 상관을 이해하기 위한 기초연구로서 1세대와 2세대로 나뉜 다섯 개의 SSD와 Ext3와 XFS 두 개의 파일 시스템에서 순차 읽기, 순차 쓰기, 그리고 임의 쓰기 워크로드에 대한 분석을 IOzone 벤치마크로 O\_DIRECT를 끈 상태에서 실험하였다. Nilfs2를 제외한 모든 파일 시스템의 성능이 일정하지 않다.

### 4. Acknowledgement

본 연구는 지식경제부 및 한국산업기술평가관리원의 산업원천기술개발사업(정보통신)의 일환으로 수행하였음 [No. 10035202, 대용량 MLC SSD 핵심기술 개발]

### 5. 참고 문헌

- [1] S. Boboila P. Desnoyers, "Write Endurance in Flash Drives: Measurements and Analysis," 2010, pp. 9-9.
- [2] F. Chen, D. A. Koufaty, X. Zhang, "Understanding Intrinsic Characteristics and System Implications of Flash Memory Based Solid State Drives," 2009, pp. 181-192.
- [3] L. Stein, "Stupid File Systems Are Better," 2005, pp. 5-5.
- [4] W. D. Norcott D. Capps, "Iozone Filesystem Benchmark," vol. 55.