

# SSD 성능 비교를 위한 I/O 트레이스 리플레이어 분석

이인혁\*, 이규환\*\*, 강윤석\*\*, 조용연\*\*, 김상욱<sup>§</sup>\*\*

\*한양대학교 공과대학 컴퓨터공학부

\*\*한양대학교 컴퓨터 소프트웨어 학과

e-mail : heyork1@gmail.com, {hwan9024, poyche, jyy0430, wook}@hanyang.ac.kr

## Analysis on I/O Trace Replayer for SSD performance evaluation

Inhyuk Yee\*, Kyuhwna Lee\*\*, Yoonsuk Kang\*\*, Yong-Yeon Jo\*\*, Sang-Wook Kim\*\*

\*Dept. of Computer Science & Engineering, Hanyang University

\*\*Dept. of Computer and Software, Hanyang University

### 요 약

본 논문에서는 I/O 트레이스 리플레이어들의 분석을 통해 SSD 성능 평가에 적합한 리플레이어를 찾는다. 또한 저장장치의 성능 비교를 위한 척도로써 적합한 척도를 도출한다. 실험결과, 시간 리플레이어가 I/O 트레이스 리플레이어로써 적합함을 밝히고, Q2C의 합이 성능 평가에 적합한 척도임을 보인다.

### 1. 서론

하드디스크에 비해 균일하고 빠른 접근속도, 에너지 저소비, 높은 내구성 등의 우월성을 가진 SSD는 모바일 기기뿐만 아니라 서버와 개인용 컴퓨터의 저장장치로 하드디스크를 대체하고 있다 [1][2].

SSD는 크게 SSD 컨트롤러, 채널, 플래시 메모리 셀 등으로 구성되어 있다. 이러한 구조적 차이로 인해 SSD의 성능이 달라질 수 있다 [1]. 이는 SSD의 I/O 접근 패턴마다 다른 성능을 보이는 원인으로 이어진다.

SSD 내부구조는 제조사마다 다르며 이에 대한 자세한 설명은 공개하지 않는다. 따라서 사용자마다 사용하는 응용프로그램이 다를 수 있고, 이마다 I/O 접근 패턴이 다르기 때문에 SSD마다 성능이 다를 수 있다. 따라서 사용자에게 적합한 SSD를 잘 고르기 위해서 I/O 접근 패턴에 대한 SSD별 성능 비교가 필요하다.

SSD는 다양한 벤치마크를 통해 성능을 평가할 수 있다. 그러나 벤치마크를 실행시키기 위해서는 환경을 구축하는데 많은 시간을 요구하게 된다. 이에 대한 대안으로 I/O 트레이스 리플레이어가 있다.

I/O 트레이스 리플레이어는 I/O 트레이스만을 통해 저장장치의 성능을 평가하는 방법이다. 이 때, I/O 트레이스란 벤치마크가 저장장치에 I/O 활동을 할 때, 각 요청이 언제, 어느 주소에, 몇 개의 블록에 대해서 어떤 작업을 했는지 등의 워크로드 정보이다. 대표적인 I/O 트레이스 리플레이어로써 (1) 시간 리플레이어와 (2) 아웃스탠딩 리플레이어가 있다.

기존에는 SSD의 성능을 비교하는데 있어 단순히 완료시간만을 측정하였다. 그러나, SSD의 성능에 영향을 미칠 만큼 충분한 부하를 주지 못하는 I/O 트레이스라면 SSD의 성능이 다르더라도 모두 동일한 수행시간을 가질 수 있다.

본 논문에서는 I/O 트레이스 리플레이어로써 적합한 리플레이어와 성능 평가 척도를 찾고자 한다. 실험 결과, 시간 리플레이어가 적합한 I/O 트레이스 리플레이어임을 보였고, 하나의 요청의 시작 시각부터 완료되는 시각까지의 시간(Q2C)의 합이 적합한 성능 비교 척도임을 보였다.

### 2. 관련연구

#### 2.1. 시간 리플레이어

시간 리플레이어는 I/O 트레이스의 각 요청이 보내진 시각을 맞춰서 I/O 트레이스를 재현한 툴이다. 일반적인 시스템 콜을 사용해서 I/O 요청을 보내면 저장장치로부터 응답을 받을 때까지 기다린 후 다음 요청을 보낼 수 있다. 만약 벤치마크가 단일 쓰레드가 아닌 멀티 쓰레드로 디바이스에 I/O 활동을 했다면 일반적인 시스템 콜을 사용해서는 벤치마크 I/O 트레이스의 요청을 보낸 시각을 맞출 수 없을 수 있다. 이 문제를 해결하기 위해 비동기 I/O를 사용한다.

#### 2.2. 아웃스탠딩(outstanding) 리플레이어

아웃스탠딩 블록 개수란 저장장치에서 대기하거나 처리 중인 블록의 개수이다. 아웃스탠딩 리플레이어는 I/O 트레이스의 각 요청을 보내기 바로 전의 아웃스탠딩 블록 개수를 유지하며 I/O 트레이스를 재현한 툴이다 [3].

본 연구는 (1) SK 하이닉스(주)의 위탁에 따른 "연구용역"과 (2) 미래창조과학부 및 정보통신산업진흥원의 대학 IT 연구센터 지원사업 (NIPA-2013-H0301-13-4009)의 연구 결과로 수행되었음.

<sup>§</sup> 교신저자

### 3. 실험

본 실험은 다음과 같은 환경에서 수행하였다. 서버는 Intel i7-4770K 3.50GHZ 프로세서와 32GB 메인 메모리로 구성되어 있다. 운영체제는 Ubuntu 를 사용하였다. 운영체제는 실험에 사용되지 않는 별도의 SSD 에 설치하였다. 이는 실험에 사용되는 SSD 가 다른 I/O 의 영향을 받지 않기 위함이다. I/O 스케줄러는 SSD 에 가장 적합한 noop I/O 스케줄러를 사용하였다. 실험에 사용한 SSD 는 잘 알려진 두 개의 SSD (SSD A 와 SSD B)이고 모두 120GB 이며, raw device 로 설정하여 실험을 수행하였다.

본 실험에서는 블록 디바이스 레벨에서 I/O 트레이스를 추출 및 분석을 위해 blktrace 와 btt 를 사용하였다 [3]. 사용한 데이터는 다음과 같다. WebSearch 는 Storage Performance Council (SPC)에서 제공한 검색엔진 I/O 트레이스이다 [4]. 원본의 완료시간은 128 초이다. TPC-C 는 해머오라[5]를 사용해 SSD A 에서 추출한 I/O 트레이스이다. 원본의 완료시간은 396 초 이다.

두 개의 성능 비교 척도를 사용한다. 하나는 I/O 트레이스의 최종 I/O 의 완료 시간 (완료시간)이다. 또 다른 하나는 하나의 요청이 시작 시각부터 완료되는 시각까지의 시간 (Q2C)의 합을 사용하였다. Q2C 는 각 I/O 의 처리되는 시간을 나타내기 때문에, SSD 의 성능 평가 척도로 볼 수 있다.

#### 3.1. 시간 리플레이어

표 1 은 시간 리플레이어를 통해 각 I/O 트레이스를 재현한 결과이다.

WebSearch 의 경우, SSD A 와 B 의 최종 완료시간이 각각 711 초, 704 초이다. 원본을 처리한 시스템의 저장장치보다 본 실험에서 사용된 저장장치의 성능이 떨어진 것으로 판단된다. 완료시간만을 통해 SSD 의 성능을 비교한 결과, SSD B 가 A 에 비해 WebSearch 에 더욱 적합한 SSD 임을 알 수 있다. 또한 Q2C 총합을 통해서도 SSD B 가 더욱 적합한 SSD 임을 보인다.

TPC-C 의 경우, SSD A 와 B 의 최종완료시간이 396 초로 모두 같다. 이러한 경우, 두 SSD 의 최종완료시간으로 성능비교가 불가능하다. 그러나 Q2C 합은 SSD A 가 1,029 초, SSD B 가 1,481 초이다. 이를 통해 SSD A 가 TPC-C 에 더욱 적합한 SSD 라고 볼 수 있다.

본 실험에서 완료시간만을 통해 SSD 의 성능을 분석하기 어렵다는 점을 알 수 있다. TPC-C 는 I/O 실행시간에 비해 대기 시간이 매우 많은 I/O 트레이스이다. 따라서 SSD 의 성능에 영향을 미칠 만큼의 충분한 부하를 주지 못하고, 대기 시간에 의해 동일한 완료시간을 갖는다. 즉, Q2C 의 총합과 같은 새로운 척도가 필요함을 도출하였고, Q2C 의 총합은 적절한 척도로 판단된다.

<표 1> 시간 리플레이어를 이용한 I/O 트레이스 재현 결과

	SSD	Q2C 합 (초)	완료시간 (초)
WebSearch	SSD A	158,118	711
	SSD B	155,705	704
TPC-C	SSD A	1,029	396
	SSD B	1,481	396

#### 3.2. 아웃스탠딩 리플레이어

표 2 는 아웃스탠딩 리플레이어를 통해 각 I/O 트레이스를 재현한 결과이다. WebSearch 의 경우, 표 1 에서 WebSearch 결과와 비슷한 결과를 보였다. 반면 TPC-C 의 경우, 표 1 에서 TPC-C 결과와 매우 다른 결과를 보였다. 트레이스에는 I/O 수행시간뿐만 아니라 대기 시간도 존재한다. 아웃스탠딩 리플레이어는 대기 시간을 무시한 채, 아웃스탠딩 블록 개수만 유지하며 I/O 를 처리한다. 따라서 표의 결과와 같이 실제 수행시간과 달리 매우 단축된 시간을 갖게 된다.

이를 통해 아웃스탠딩 리플레이어가 I/O 트레이스의 워크로드를 제대로 재현하지 못한다고 볼 수 있다. 따라서 아웃스탠딩 리플레이어는 I/O 트레이스 리플레이어로 적합하지 못하다.

<표 2> 아웃스탠딩 리플레이어를 이용한 I/O 트레이스 재현 결과

	SSD	Q2C 합 (초)	완료시간 (초)
WebSearch	SSD A	160,181	707
	SSD B	152,211	704
TPC-C	SSD A	1,259	54
	SSD B	1,747	146

### 4. 결론

본 논문에서는 시간 리플레이어와 아웃스탠딩 리플레이어가 I/O 트레이스 리플레이어으로써 적합한지 판단하고 각 리플레이어를 수행 후, Q2C 의 합을 척도로 두 SSD 의 성능을 비교해보았다. 시간 리플레이어로 재현했을 경우 원본 I/O 트레이스의 워크로드를 잘 재현한 것을 확인하였다. 그러나 아웃스탠딩 리플레이어로 충분한 부하를 주지 못하는 트레이스를 재현했을 경우, 완료시간이 원본 트레이스보다 작게 나온다. 이를 통해 아웃스탠딩 리플레이어가 I/O 트레이스 리플레이어으로써 부적합한 것을 확인했다. 또한, 새로운 성능비교 척도인 Q2C 의 합으로 시간 리플레이어로 I/O 트레이스를 재현한 후 두 SSD 의 성능을 비교할 수 있었다.

#### 참고문헌

- [1] Feng Chen, Rubao Lee, and Xiaodong Zhang, "Essential Roles of Exploiting Internal Parallelism of Flash Memory based Solid State Drives in High-Speed Data Processing," In *Proc. IEEE Int'l Symp. on High Performance Computer*, IEEE HPCA, pp. 266-277, 2011.
- [2] Yang Hu, Hong Jiang, Dan Feng, Lei Tian, Hao Luo, and Chao Ren, "Exploring and Exploiting the Multi-level Parallelism Inside SSDs for Improved Performance and Endurance," *IEEE Trans. on Computers*, Vol. 62, No. 6, pp 1141-1155, 2013.
- [3] Sankaran Sivathanu, Jinpyo Kim, Devaki Kulkarni, Ling Liu, "Load-Aware Replay of I/O Traces," In *Proc. of Int'l USENIX Conf. on File and Storage Technologies*, FAST, 2011.
- [4] Blktrace & Btt [Internet], <http://linux.die.net/man/>
- [5] HammerDB [Internet], <http://hammerora.sourceforge.net>