

낸드 플래시 메모리 저장장치에서 읽기 성능 향상을 위한 분석 및 설계

송수미¹, 이현섭^{2*}

*교신저자

¹백석대학교 컴퓨터공학부 학부생

²백석대학교 컴퓨터공학부 교수

mjsong2001@bu.ac.kr, hyunseob@bu.ac.kr

Analysis and Design for Improving Read Performance on NAND Flash Memory Storage

Su-Mi Song¹, Hyun-Seob Lee²

^{1,2}Dept. of Computer Engineering, Baekseok University

요 약

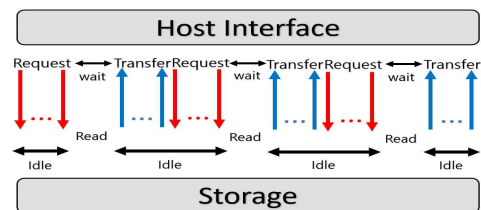
최근 빅데이터를 처리하기 위한 고용량의 저장장치의 연구가 활발하게 진행되고 있다. 특히 낸드 플래시 메모리 기반 저장장치인 SSD(Solid State Drive)는 고속 데이터 처리가 가능하기 때문에 다양한 저장 시스템에서 사용되고 있다. 그러나 낸드 플래시 메모리는 읽기 속도가 빠르고 쓰기 속도가 느린 비대칭 구조를 가지고 있어 데이터 쓰기 작업이 진행 중일 때 다른 작업이 대기하게 되어 전체적인 시스템 성능에 영향을 미칠 수 있다. 이에 대해 대부분의 연구에서는 쓰기 성능을 개선하기 위한 버퍼 관리 정책들이 제안되었다. 기술의 발전으로 쓰기 성능은 개선되었지만, 그에 비해 읽기 성능이 저조한 문제가 있다. 본 논문에서는 읽기 성능이 쓰기 성능보다 취약한 점을 분석하고 플래시 메모리 내에 저장장치 속도를 개선하기 위해 앞으로 읽기 요청 가능성이 있는 예상 데이터를 선형회귀 모델을 적용하여 전송 대기 시간 중 미리 읽는 정책을 제안한다.

1. 서론

최근 대용량 데이터 처리를 위한 고성능의 저장장치에 대한 필요성이 증가하면서 대부분의 엔터프라이즈에서는 대용량의 데이터를 처리하기 위해 빠른 입출력 성능을 가진 낸드 플래시 메모리 기반 저장장치인 SSD(Solid State Drive)를 사용하고 있다. SSD 내부에서 데이터를 저장하는 공간인 낸드 플래시 메모리의 읽기와 쓰기 수행은 페이지 단위로 접근할 수 있고 이미 데이터가 존재하는 페이지에는 덮어쓰기가 불가능하며 쓰기 연산을 하기 위해서는 소거 연산을 수행 해야한다. 이러한 플래시 메모리의 구조적 한계를 감추기 위해 FTL(Flash Translation Layer)을 사용한다.[1] 또한 플래시 메모리는 읽기 속도에 비해 쓰기 속도가 느린 특징을 가지고 있다. 최근까지 버퍼 관리 정책은 읽기 속도가 느려지더라도 쓰기 작업을 효율적으로 처리할 수 있도록 설계되었다. 이로 인해 쓰기 성능이 개선되었지만 최근 SSD 내부에서는 쓰기 성능보다 읽기 성능이 취약한 점을 보이고 있다. 본 논문에서는 읽

기 성능 문제를 해결하기 위해 선형회귀 모델을 이용한 LRU(Least Recently Used)[2] 기반 버퍼 관리 정책을 제안한다.

2. 관련 연구



(그림 1) 저장장치의 데이터 처리 과정

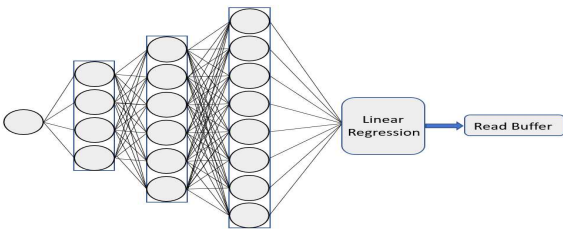
그림1은 호스트와 저장장치 사이의 분리된 데이터 전송 영역과 데이터 처리 영역의 구조를 보여주고 있다. 호스트로부터 데이터 읽기 요청을 받아서 수행하고 저장 장치로부터 읽기 처리된 데이터를 호스트로 전송하는 구조로 되어있다. 호스트에서 읽기 요청된 주소를 저장 장치로부터 데이터를 읽은 후 버퍼에 모아서 호스트로 전달할 수 있도록 한다. 데이터 처리 영역은 저장장치와 하드웨어 모듈로 분리

되어 있기 때문에 데이터 입출력이 병렬로 수행될 수 있다. 그러나 동기화된 데이터로 수행한다면 연속적인 데이터 읽기 처리가 수행되는 경우 그림1과 같이 호스트는 다음 읽기 처리를 수행하기 위해 저장장치로부터 데이터 전송을 기다리게 되고, 저장장치는 이전 데이터 전송이 완료된 후에 다음 요청을 받아 처리할 수 있다. 이러한 상황에서 호스트와 저장장치 간에 데이터 이동 속도보다 데이터가 내부적으로 처리되는 속도가 빠르기 때문에 다음 읽기 처리를 수행하기 전까지 유휴 시간이 발생할 수 있다.[3]

3 데이터 예측 정책

3.1 선형회귀 기반 분석기 실험 설계

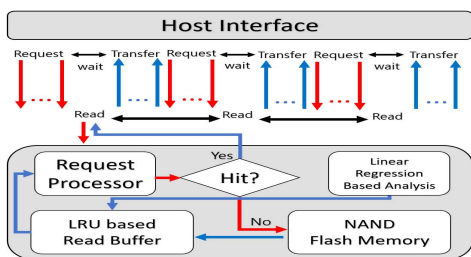
그림2는 본 논문에서 제안하는 분석기의 내부 구조이다. 제안하는 예상 데이터 미리 읽기 정책은 선형회귀를 기반으로 다음 읽기 요청을 예측한다. 선형회귀(linear regression)는 독립변수와 입력 변수 사이의 관계를 모델링 하는 기법으로 데이터의 값을 분석하는 방법이다.



(그림 2) 분석기 내부 구조

신경망 구조를 통해 호스트로부터 읽기 요청된 데이터의 패턴을 분석하여 다음 요청될 데이터를 예측하도록 한다. 선형회귀 모델은 학습된 신경망의 출력 값을 입력으로 받아 선형회귀를 수행하고 다음 주소 데이터 예측 결과를 출력한다. 예측된 주소 데이터의 값은 버퍼로 이동시켜 다음 읽기 요청에 대비할 수 있도록 한다.

3.2 분석기를 적용한 저장장치 구조



(그림 3) 분석기를 적용한 데이터 예측 과정

그림3의 저장장치의 내부 구조는 본 논문에서 제안하는 데이터 예측 분석기를 포함한 그림이다. 호스트로부터 읽기 요청을 받아 처리할 때 요청된 데이터를 호스트로 전달하는 동안 분석기에서 데이터 접근 패턴을 분석하여 예상되는 다음 읽기 요청 데이터를 선별한다. 선별된 예측 데이터를 미리 버퍼에 올려두었을 때, 다음 요청된 주소의 데이터가 올려 둔 데이터와 일치하는 경우 버퍼에서 데이터를 가져와 처리할 수 있기 때문에 읽기 연산을 수행할 때 메모리에서 데이터를 가져오는 것보다 빠른 속도로 처리할 수 있다. 이러한 방식은 낸드 플래시 메모리로부터 불필요한 읽기 횟수를 줄일 수 있는 반면에 일치하지 않는 경우 메모리로부터 데이터를 읽어서 호스트에 전달된다.

4 결론

본 논문에서는 플래시 메모리 내에 저장장치 속도를 개선하기 위해 앞으로 읽기 요청 가능성이 있는 예상 데이터를 선형 회귀 모델을 적용하여 전송 대기 시간 중 미리 읽는 정책을 제안하였다. 이 방법을 통해 읽기 대기 시간을 최적화하여 SSD의 읽기 성능을 개선할 수 있다. 추후 논문에서는 제안하는 예측 데이터 미리 읽기 정책을 적용하여 플래시 메모리 기반 SSD에서 읽기 성능을 비교하는 연구를 진행할 예정이다.

Acknowledgement

이 논문은 2023년도 정부(교육부)의 재원으로 한국 연구재단의 지원을 받아 수행된 기초연구사업임 (NRF2021R1I1A3061020)

참고문헌

[1] Seung-woo Lee, Se-Jin Oh, "Index management technique using Small block in storage device based on NAND flash memory" Journal of The Korea Society of Computer and Information, Vol. 25 No. 10, pp. 1-14

[2] Seon Kim, Sungjae Lee, Inhwan Lee, ".Buffer Cache Block Replacement Algorithm Based on LRU and Prefetching", 한국정보과학회 학술발표논문집, 37(2B), 298-303. (2010)

[3] H.-S. Lee, "A Prediction-Based Data Read Ahead Policy using Decision Tree for improving the performance of NAND flash memory based storage devices," Journal of Internet of Things and Convergence, vol. 8, no. 4, pp. 9 - 15, Aug. 2022.