

MLC 낸드 플래시 기반 저장장치의 쓰기 성능 개선을 위한 계층 교차적 최적화 기법

(A Cross Layer Optimization Technique for Improving Performance of MLC NAND Flash-Based Storages)

박 지 성 ^{*} 이 성 진 ^{**} 김 지 흥 ^{***}
(Jisung Park) (Sungjin Lee) (Jihong Kim)

요 약 하나의 메모리 셀에 여러 비트의 정보를 저장하는 다치화 기법은 공정 미세화와 함께 낸드 플래시 메모리의 집적도를 크게 향상시켰지만, 그 반대급부로 MLC 낸드 플래시 메모리의 평균 쓰기 성능은 SLC 낸드 플래시 메모리 대비 두 배 이상 하락하였다. 본 논문에서는 MLC 낸드 플래시 기반 저장장치의 성능 향상을 위해 제안되었던 기존의 계층 교차적 최적화 기법들을 소개하고, 두 기법의 상호 보완성을 분석하여 해당 기법들의 한계점을 극복하는 새로운 통합 기법을 제안한다. MLC 낸드 플래시 디바이스에 존재하는 성능 비대칭성을 플래시 변환 계층 수준에서 최대한 활용함으로써, 제안하는 기법은 인가되는 다수의 쓰기 명령을 SLC 낸드 플래시 디바이스의 성능으로 처리하여 저장장치의 성능 향상을 도모한다. 실험 결과, 제안하는 기법은 기존 기법 대비 평균 39%의 성능 향상을 달성할 수 있음을 확인하였다.

키워드: 멀티 레벨 셀 낸드 플래시 메모리, 낸드 플래시 기반 저장장치, 플래시 변환 계층, 계층 교차적 최적화

Abstract The multi-leveling technique that stores multiple bits in a single memory cell has significantly improved the density of NAND flash memory along with shrinking processes. However, because of the side effects of the multi-leveling technique, the average write performance of MLC NAND flash memory is degraded more than twice that of SLC NAND flash memory. In this paper, we introduce existing cross-layer optimization techniques proposed to improve the performance of MLC NAND flash-based storages, and propose a new integration technique that overcomes the limitations of existing techniques by exploiting their complementarity. By fully exploiting the performance asymmetry in MLC NAND flash devices at the flash translation layer, the proposed technique can handle many write requests with the performance of SLC NAND flash devices, thus significantly improving the performance of NAND flash-based storages. Experimental results show that the proposed technique improves performance 39% on average over individual techniques.

Keywords: multi-level cell NAND flash memory, NAND flash-based storages, flash translation layer (FTL), cross-layer optimization

- 이 연구를 위해 연구장비를 지원하고 공간을 제공한 서울대학교 컴퓨터연구소에 감사드립니다.
- 이 논문은 2017 한국연구재단의 지원으로 서울대학교 컴퓨터공학부 BK21플러스 컴퓨터미래인재양성사업단의 지원을 받아 수행된 연구임 (21A20151113068)
- 이 논문은 2016년도 정부(미래창조과학부)의 지원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2015M3C4A7065645)
- 이 논문은 제43회 동계학술발표회에서 'MLC 낸드 플래시 기반 저장장치의 쓰기 성능 개선을 위한 계층 통합적 최적화 기법'의 제목으로 발표된 논문을 확장한 것임

- ^{*} 학생회원 : 서울대학교 컴퓨터공학부
jspark@davinci.snu.ac.kr
- ^{**} 정 회 원 : 대구경북과학기술원 정보통신융합전공 교수
sungjin.lee@dgist.ac.kr
- ^{***} 종신회원 : 서울대학교 컴퓨터공학부 교수(Seoul Nat'l Univ.)
jihong@davinci.snu.ac.kr
(Corresponding author임)

- 논문접수 : 2017년 2월 20일
(Received 20 February 2017)
- 논문수정 : 2017년 7월 23일
(Revised 23 July 2017)
- 심사완료 : 2017년 7월 24일
(Accepted 24 July 2017)

Copyright©2017 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.
정보과학회논문지 제44권 제11호(2017. 11)

1. 서론

낸드 플래시 메모리는 높은 내충격성(impact resistance)과 저전력의 장점으로 인해 전통적으로 모바일 시스템에서 필수적으로 차용되어왔으나, 차세대 컴퓨팅 환경의 저장장치에 대한 높은 성능 요구조건으로 인해 엔터프라이즈 서버와 같은 영역에서도 하드디스크를 빠르게 대체하고 있다. 특히, 하나의 메모리 셀에 여러 비트의 정보를 저장하는 다치화 기법(multi-leveling technique)은 반도체 공정의 미세화와 함께 낸드 플래시 메모리의 집적도를 크게 높여, 낸드 플래시 기반 저장장치가 다양한 컴퓨팅 환경에서 폭 넓게 활용되는 데 크게 기여하였다.

다치화 기술은 낸드 플래시 기반 저장장치의 단위 용량 당 가격을 크게 감소시켰으나, 그 반대급부로 기존 낸드 플래시 메모리와 비교하여 심각한 쓰기 성능의 하락을 초래하였다. 예를 들어, 하나의 메모리 셀에 두 비트의 정보를 기록하는 2-bit MLC(Multi-Level Cell) 낸드 플래시 메모리는 하나의 메모리 셀에 한 비트만을 기록하는 SLC(Single-Level Cell) 낸드 플래시 메모리에 비해 쓰기 명령을 처리하는 데 두 배 이상의 시간을 소모한다. 이는 MLC 낸드 플래시 메모리의 기록 방식에 기인한 것으로, 20 나노미터 공정의 낸드 플래시 메모리를 예로 들면, 처음 저장되는 비트(Least-Significant Bit, LSB)를 저장할 때는 SLC 낸드 플래시 메모리와 비슷한 500 us의 시간이 소모되지만, 두 번째 비트(Most-Significant Bit, MSB)를 저장할 때에는 더욱 미세한 전력 제어가 필요하여 네 배에 가까운 2000 us의 시간이 소모된다[1]. 빅데이터 시스템 및 클라우드 기반 컴퓨팅과 같은 차세대 컴퓨팅 환경에서 저장장치에 대한 성능 요구조건이 지속적으로 높아짐에 따라, 이와 같은 MLC 낸드 플래시 기반 저장장치의 성능 하락은 반드시 해결해야 할 문제로 주목받고 있다.

MLC 낸드 플래시 메모리의 성능 향상을 위해, 디바이스 계층의 특성을 낸드 플래시 기반 저장장치의 펌웨어 계층인 FTL(Flash Translation Layer, 플래시 변환 계층)에서 적극 활용하는 계층 교차적 최적화 연구들이 진행된 바 있다. 예를 들어, MLC 낸드 플래시 메모리 상에 존재하는 불필요한 페이지 기록 순서를 완화하여 높은 성능이 요구될 때 빠른 LSB 페이지를 집중적으로 사용함으로써 실질적인 체감 성능을 향상시키는 Relaxed Program Sequence(RPS) 기법이 제안된 바 있으며[2], 메모리 셀을 공유하는 LSB 페이지의 데이터가 더 이상 유효하지 않을 때, 페이지 기록 방식을 바꿔 MSB 페이지를 LSB 페이지에 준하는 성능으로 기록하는 SLC-Like Program(SLP) 기법 또한 제안되었다[3]. 해당 기법들은 기존 연구들에서 활용하지 못한 MLC 낸드 플래시

디바이스의 성능 비대칭성을 상위 소프트웨어 계층에서 적극 활용하는 디바이스-펌웨어 간 계층 교차적 최적화를 통해 큰 하드웨어의 수정 없이도 높은 성능 개선을 달성할 수 있었다.

위와 같은 기존 계층 교차적 최적화 기법들은 특정 조건의 작업부하에서는 매우 높은 성능 개선을 달성할 수 있으나, 해당 조건들이 만족하지 않을 경우에는 성능 개선을 거의 기대할 수 없는 한계를 보인다. 예를 들어, RPS 기법은 빠른 LSB 페이지를 연속적으로 사용하여 순간 성능을 높일 수 있지만, 쓰기 명령이 유희 시간 없이 지속적으로 집중된다면 일정량의 LSB 페이지를 소모한 후 MSB 페이지를 사용해야 하므로 기존 MLC 낸드 플래시 기반 저장장치의 성능과 동일한 성능을 보이게 된다. SLP 기법은 MSB 페이지를 빠르게 기록하기 위해서 대응되는 LSB 페이지의 데이터가 무효화(invalidation)돼야 하는데, 기존 페이지 기록 순서를 따르게 된다면 LSB 페이지와 MSB 페이지를 번갈아 사용해야 하므로 LSB 페이지 데이터의 무효화 확률이 매우 낮고, 따라서 최적화 기법이 거의 적용되지 못하게 된다.

본 논문에서는 기존에 제안된 RPS 기법과 SLP 기법이 상호 보완적이라는 고찰을 통해, 두 기법들의 한계점을 극복할 수 있는 새로운 통합 기법을 제안한다. 새로 제안하는 기법이 적용된 낸드 플래시 기반 저장장치에서는, 인가되는 쓰기 명령을 (RPS 기법을 적용하여) LSB 페이지를 우선적으로 집중 사용하여 빠르게 처리하고, 이후 해당 LSB 페이지의 데이터가 무효화 되었을 때 나머지 MSB 페이지도 (SLP 기법을 적용하여) 빠르게 기록함으로써 MLC 낸드 플래시 기반 저장장치의 성능을 SLC 낸드 플래시 기반 저장장치에 최대한 가깝게 향상시킨다. 시뮬레이션을 통한 실험결과, 제안하는 기법이 적용된 저장장치는 대부분의 작업 부하에 대해 빠른 페이지 쓰기를 통해 처리되는 쓰기 명령의 비율을 크게 향상시킬 수 있었으며, 이를 통해 평균 39%의 대역폭 향상을 달성할 수 있음을 확인하였다.

이어지는 본 논문의 구성은 다음과 같다. 0절에서는 배경지식 및 선행연구로서 MLC 낸드 플래시의 특성과 통합의 대상이 되는 기법들을 설명한다. 3절에서는 두 기법의 통합 가능성에 대해 설명하고, 새로운 통합 최적화 기법을 제안한다. 4절에서는 시뮬레이션을 통한 실험 결과를 보이고, 5절에서 결론을 맺는다.

2. 배경지식 및 선행연구

2.1 MLC 낸드 플래시 메모리의 기록 방식

낸드 플래시 메모리는 메모리 셀의 임계 전압(threshold voltage)을 조정하여 데이터를 기록한다. 메모리 셀의 임계 전압은 플로팅 게이트(floating gate)에 존재

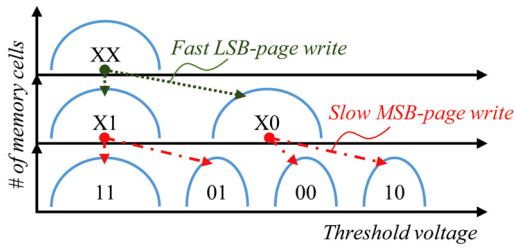


그림 1 MLC 낸드 플래시 메모리의 기록 방식
 Fig. 1 An illustration of the program scheme for MLC NAND flash memory

하는 전자의 양에 따라 달라지고, 플로팅 게이트에 존재하는 전자의 양은 전자를 추출하는 삭제 동작이나 전자를 삽입하는 기록 동작이 발생하지 않는 한 오랜 시간 유지되므로, 낸드 플래시 메모리는 비휘발성 메모리로 사용될 수 있다. 그림 1은 MLC 낸드 플래시 메모리의 기록 방식을 보여준다. 모든 메모리 셀의 임계 전압을 정확히 같게 맞출 수 없기 때문에, 임계 전압은 그림 1과 같이 분포로써 설정되는데, MLC 낸드 플래시 메모리는 더 많은 임계 전압 분포를 만들고 각 분포에 데이터를 사상함으로써 같은 메모리 셀에 더 많은 비트를 저장할 수 있다. 그림 1에서 보이듯, MLC 낸드 플래시 메모리는 하위 비트 정보를 담은 LSB 페이지를 위한 두 개의 분포를 먼저 생성하고, 이를 기반으로 상위 비트 정보를 담은 MSB 페이지를 위한 최종적인 네 개의 임계전압 분포를 생성한다.

한편, 임계 전압 분포가 존재할 수 있는 전압 범위 (voltage window)는 일정하므로, 제한된 전압 범위에 더 많은 수의 임계 전압 분포를 설정하기 위해서는 분포의 폭이 좁아야 하고, 이는 플로팅 게이트에서 전자를 삽입할 때 더욱 세밀한 제어가 필요하여 데이터 기록 시간이 증가함을 의미한다. 그림 1에서와 같이, LSB 페이지의 기록 시에는 두 개의 임계 전압 분포만을 생성하기 때문에 SLC 낸드 플래시 메모리에 준하는 성능을 보이지만, MSB 페이지의 기록 시에는 네 개의 임계 전압 분포를 생성해야 하므로 SLC 낸드 플래시 메모리의 페이지 기록보다 네 배 긴 시간이 요구된다. 이와 같이 MLC 낸드 플래시 메모리에서는 페이지 간 큰 성능 비대칭성이 존재하며, 결국 MLC 낸드 플래시 메모리의 쓰기 성능은 LSB 페이지와 MSB 페이지의 평균 쓰기 성능으로 표현할 수밖에 없고, MSB 페이지의 낮은 쓰기 성능으로 인해 SLC 낸드 플래시 메모리 보다 두 배 이상의 악화된 평균 쓰기 성능을 보이게 된다.

2.2 Relaxed Program Sequence (RPS) 기법

MLC 낸드 플래시 메모리 상에 존재하는 성능 비대

칭성을 상위 계층인 FTL 수준에서 활용하면 낸드 플래시 기반 저장장치의 체감 성능을 크게 향상시킬 수 있다. 예를 들어, 쓰기 명령이 집중돼서 인가될 경우에는 빠른 LSB 페이지를 집중적으로 사용하여 쓰기 명령의 지연시간을 최소화하고, 적은 쓰기 명령이 간헐적으로 인가될 때나 유휴 시간 중의 가비지 컬렉션 때는 느린 MSB 페이지를 사용하면 같은 디바이스를 사용하면서도 체감 성능을 높일 수 있다.

그럼에도 불구하고, 실제 MLC 낸드 플래시 메모리에서는 강력한 페이지 기록 순서 제약으로 인해 위와 같은 성능 비대칭성을 상위 계층에서 활용하는 것이 불가능하다. 이는 MLC 낸드 플래시 메모리에서 데이터 신뢰성을 보장하기 위함으로, 인접한 페이지 간의 셀 간 간섭(cell-to-cell interference)를 최소화하기 위함이다 [4]. 셀 간 간섭을 최소화하기 위해서는 하나의 워드 라인(word line: 페이지가 저장되는 메모리 셀의 집합)에 LSB 페이지와 MSB 페이지가 모두 저장된 후, 인접한 워드 라인의 페이지 쓰기 횟수를 최소화해야 하는데, 이를 위해 메모리 제조사에서는 LSB 페이지와 MSB 페이지를 번갈아 사용하게 되는 블록 내 페이지 기록 순서를 엄격히 따를 것을 요구한다. 그림 2(a)는 일반적인 MLC 낸드 플래시 메모리에서의 블록 내 페이지 기록 순서를 보여준다. 그림 2(a)에서 확인할 수 있듯, 모든 워드 라인은 MSB 페이지가 기록된 후 한 번의 간섭(바로 위 워드 라인의 MSB 페이지 쓰기)만을 받게 된다. 이와 같은 페이지 기록 순서를 따른다면, 성능 요구 조건이 높을 때 빠른 LSB 페이지를 연속적으로 사용할 수 없으므로, 성능 비대칭성을 활용한 성능 개선은 불가능하다.

RPS 기법은 위와 같은 기존 블록 내 페이지 기록 순서에 과도한 제약 사항이 존재한다는 분석 결과를 바탕으로, 불필요한 제약 사항의 완화를 통해 블록 내 LSB 페이지의 연속적인 사용을 가능하게 한 기법이다. 그림 2(b)는 불필요한 제약 사항이 완화된 블록 내 페이지 기록 순

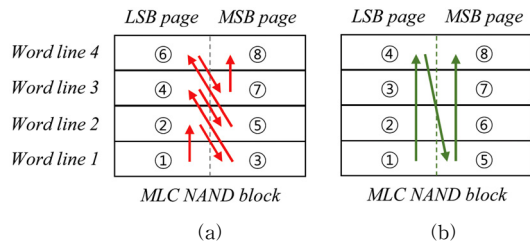


그림 2 (a) 기존 블록 내 페이지 기록 방식과 (b) RPS 기법과의 비교

Fig. 2 A comparison of (a) the conventional program sequence and (b) Relaxed Program Sequence (RPS) scheme

서를 보여준다. 그림 2(b)에서 확인할 수 있듯, 기존 페이지 기록 순서와 마찬가지로 RPS 기법 하에서도 모든 워드 라인이 MSB 페이지가 기록된 후 최대 한 번의 간섭만을 받게 된다. 결국, 블록 내에서 보다 유동적인 페이지 사용이 가능하게 되고, 요구 성능에 따른 페이지 할당을 통해 저장장치의 체감 성능 개선을 달성할 수 있다.

RPS 기법이 적용된 낸드 플래시 기반 저장장치는 쓰기 명령이 일시적으로 집중되는 환경에서는 큰 체감 성능의 향상을 보이지만, 쓰기 명령이 지속적으로 집중될 때는 성능 개선을 달성하기 어려운 한계를 보인다. 만약 지속적으로 집중되는 쓰기 명령을 처리하기 위해 제한 없이 LSB 페이지를 사용한다면, 모든 LSB 페이지의 소진 후에는 느린 MSB 페이지만을 사용하여 쓰기 명령을 처리해야 하므로 해당 시점 이후에는 저장장치의 심각한 성능 하락이 발생하게 된다[2]. 이를 막기 위해 일정 LSB 페이지의 소진 후에는 MSB 페이지를 번갈아 사용해야 하고, 결국 기존 페이지 기록 순서가 적용된 저장장치와 같은 성능을 보이게 된다.

2.3 SLC-Like Program (SLP) 기법

SLP 기법은 LSB 페이지 데이터가 더 이상 유효하지 않을 경우, 대응되는 MSB 페이지를 빠르게 기록할 수 있는 기법이다. 그림 3은 SLP 기법이 적용된 MLC 낸드 플래시 메모리의 페이지 기록 방식을 보여준다. 2.1절에서 설명한 바와 같이, LSB 페이지의 기록에 비해 MSB 페이지 기록은 더 많은 수의 임계 전압 분포를 생성해야 하기 때문에 더 긴 시간이 필요하다. SLP 기법은 MSB 페이지 기록 시 대응되는 LSB 페이지의 데이터가 무효화 되었다면, 하나의 임계 전압 분포만 생성하여 MSB 페이지 데이터를 기록하여도 무방하다는 사실을 활용하였다. 즉, LSB 페이지의 데이터가 더 이상 필요하지 않다면, 메모리 셀에는 MSB 페이지만을 기록해도 되고, 따라서 총 두 개의 임계 전압 분포만이 필요하게 된다. 이 때, 그림 3과 같이 기존에 LSB 페이지 데이터를 위해 설정된 임계 전압 분포들을 모두 '1'로 대응시키고, 하나의 임계 전압 분포만 추가로 생성하여 '0'으로 대응시키면 보다 넓은 폭을 갖는 하나의 임계 전압 분포의 생성만으로 MSB 페이지를 기록할 수 있다. 넓은 폭을 갖는 임계 전압 분포의 생성에는 세밀한 전압 제어가 요구되지 않으므로, 결과적으로 기존 LSB 페이지와 유사한 쓰기 성능으로 MSB 페이지를 기록할 수 있게 된다.

SLP 기법은 LSB 페이지가 무효화 된 경우 MSB 페이지 기록 시간을 획기적으로 단축시킬 수 있지만, 실제 MLC 낸드 플래시 메모리에서는 그 적용 기회가 매우 적을 수 있다. 2.2절에서 언급된 바와 같이, 일반적인 MLC 낸드 플래시에서는 블록 내에서의 페이지 기록

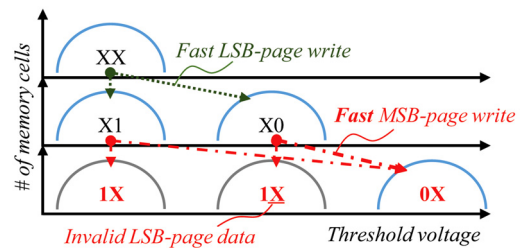


그림 3 SLP 기법의 페이지 기록 방식

Fig. 3 An illustration of SLC-Like Program (SLP) scheme

순서가 엄격히 제한되고, 그림 2(a)에서와 같이 메모리 셀을 공유하는 LSB 페이지와 MSB 페이지는 거의 연속하여 사용하게 된다. 낸드 플래시 기반 저장장치의 병렬성을 고려하더라도, 위와 같은 페이지 기록 순서 하에서는 MSB 페이지의 기록 전에 LSB 페이지의 데이터가 무효화될 가능성이 매우 낮으며, 결과적으로 SLP 기법이 빈번하게 적용되기 위해서는 작업 부하가 극도로 높은 시간적 지역성(temporal locality)을 가져야만 한다. 그러나 엔터프라이즈 서버와 같은 환경에서는 큰 단위의 쓰기 명령이 집중적으로 인가되는 경향이 있으므로, 쓰기 명령 간에 높은 시간적 지역성이 있더라도 대응되는 LSB 페이지의 데이터가 MSB 페이지의 기록 전에 무효화될 확률은 매우 낮고, 결과적으로 SLP 기법이 적용되지 않은 저장장치와 같은 성능을 보일 수밖에 없게 된다.

3. RPS 기법과 SLP 기법의 통합 최적화

3.1 기존 기법들 간의 상호 보완성

본 연구의 핵심적인 통찰은 각 기법의 개별적 적용만으로는 쓰기 명령의 집중도가 높은 작업 부하에 대해서 MLC 낸드 플래시 기반 저장장치의 성능 향상을 기대하기 어려우나, 두 기법들은 상호 보완이 가능하므로 통합을 통해 그 한계를 극복할 수 있다는 것이다. 2.2절과 2.3절에서 설명한 바와 같이, RPS 기법의 한계는 MSB 페이지 쓰기 성능 자체는 개선시키지 못한다는 것에 기인하며, SLP 기법의 한계는 대응되는 LSB 페이지와 MSB 페이지가 매우 짧은 시간 간격으로 사용되어 LSB 페이지 데이터의 무효화 확률이 낮다는 것에 기인한다. 한편, RPS 기법은 블록 내에서 LSB 페이지를 연속적으로 사용하고 향후 MSB 페이지를 사용하는 것을 가능하게 하므로 대응되는 LSB 페이지와 MSB 페이지의 사용 간 시간적 간격을 크게 늘릴 수 있으며, SLP 기법은 MSB 페이지의 쓰기 성능을 LSB 페이지에 가깝게 향상시키는 것이 가능하다. 결과적으로, 각 기법의 핵심적 기여가 서로의 한계점을 직접적으로 극복 가능하게 한다.

위와 같이 상호보완적인 두 기법을 통합하면, 다음과 같은 쓰기 성능 최적화가 가능하다. 우선 인가되는 쓰기 명령을 RPS 기법을 적용하여 LSB 페이지를 집중적으로 사용하여 처리한다. 기존 RPS 기법만이 적용된 저장 장치보다 더욱 공격적으로 LSB 페이지를 사용하여, 블록 내의 LSB 페이지와 MSB 페이지가 사용되는 시점 간 간격을 넓혀 모든 LSB 페이지를 소모한 블록 내의 데이터가 무효화 될 확률을 증가시킨다. 이후 LSB 페이지의 데이터가 무효화된 MSB 페이지를 SLP 기법을 적용하여 빠르게 기록한다. 위와 같은 방법의 페이지 할당 방식은 데이터의 업데이트 비율이 일정 수준 이상일 때 인가되는 쓰기 명령을 높은 확률로 LSB 페이지에 준하는 성능으로 기록할 수 있게 한다.

3.2 3-phase 블록 관리 기법

3.1절에서 언급된 통합 최적화를 위해, 본 논문에서는 블록을 페이지 사용 양상에 따라 세 단계로 구분하여 관리하는 3-phase 블록 관리 기법을 개발하였다. 그림 4는 제안하는 블록 관리 기법 하에서 블록의 순환 과정을 보여준다. 초기 상태의 모든 블록은 Fast 블록으로, 해당 블록들은 빠른 LSB 페이지만을 우선적으로 사용하여 쓰기 명령을 처리한다. Fast 블록 하나를 선택하여 FA블록(Fast-Active 블록)으로 지정하여 높은 성능 요구조건인 쓰기 명령을 처리하며, FA블록 내의 모든 LSB 페이지가 소모되면 Slow 블록으로 설정한 후 SBQueue (Slow Block Queue)에 삽입한다. SBQueue에서 관리되는 slow 블록 중 하나를 SA블록(Slow-Active 블록)으로 삼아 인가되는 쓰기 명령 중 성능 요구조건이 낮은

명령들을 처리한다. 새로운 SA블록을 선택할 때는, slow 블록들 중 유효한 데이터를 갖는 LSB 페이지가 가장 많은 블록을 선택한다. Slow 블록 중 SA블록으로 선택되어 MSB 페이지를 사용하기 이전에 모든 LSB 페이지가 무효화된 블록의 경우, Revived 블록으로 재 설정되어 RBQueue (Revived Block Queue)에서 관리된다. Revived 블록은 MSB 페이지만을 사용 가능하지만, 모든 LSB 페이지의 데이터가 무효화 됐으므로, SLP 기법을 적용하여 MSB 페이지 역시 LSB 페이지에 준하는 성능으로 기록할 수 있다. RBQueue에 삽입된 순서대로, revived 블록은 RA블록(Revived-Active 블록)으로 설정되어 인가되는 쓰기 명령 중 높은 성능 요구조건을 갖는 명령들을 처리한다. SA블록과 RA블록 내의 모든 MSB 페이지가 소모되면 새로운 SA블록, RA블록을 각각 설정한 후, 기존 블록들은 가비지 컬렉션의 대상이 되어 다시 fast 블록으로 관리된다.

3.3 동적 페이지 할당

제안하는 기법이 적용된 저장장치에서는 쓰기 버퍼를 이용하여 시스템의 성능 요구조건을 판단하고, 그에 따라 가장 알맞는 페이지를 선택하여 쓰기 명령을 처리한다. 이는 쓰기 명령이 간헐적으로 인가된다면, 적은 양의 모든 쓰기 명령을 DRAM 쓰기 버퍼에서 임시 저장하여 매우 빠르게 처리하는 것이 가능하므로, 미래에 집중될 수 있는 쓰기 명령을 처리하기 위한 빠른 페이지들의 낭비를 막기 위함이다. 이를 통해 상위 시스템 계층에서 체감하는 저장장치 성능을 I/O 접근 양상에 따라 최적으로 향상시킬 수 있다.

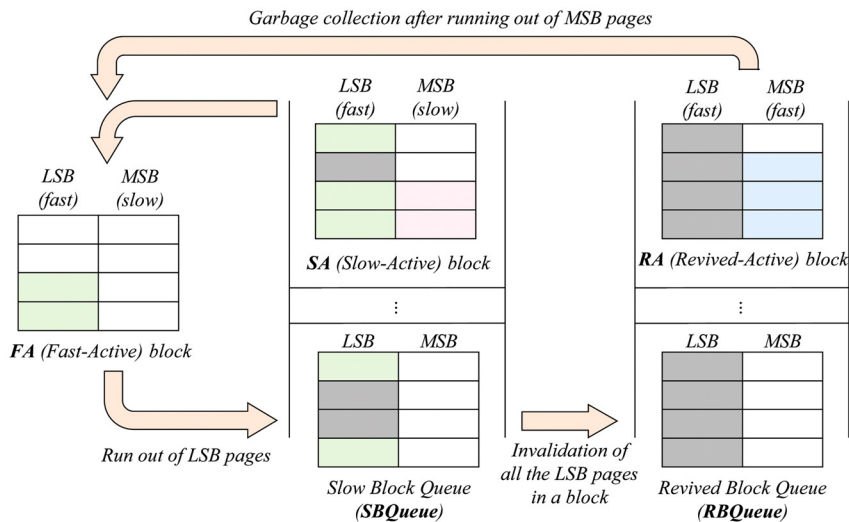


그림 4 3-phase 블록 관리 기법 하에서의 블록 순환 과정
 Fig. 4 A life cycle of a NAND block in 3-phase block management

시스템의 성능 요구조건 수준이 높을 경우, 즉, 쓰기 버퍼의 점유율이 기준치 이상(예: 80% 이상)일 경우에는 쓰기 버퍼에 임시 저장된 쓰기 명령들을 높은 쓰기 성능을 보이는 FA블록의 LSB 페이지나 RA블록의 MSB 페이지를 사용하여 처리한다. 반대로, 쓰기 버퍼의 점유율이 기준치 이하(예: 20% 이하)일 때는 SA블록의 느린 MSB 페이지만을 사용하여 쓰기 명령을 처리한다. 중간 수준의 성능 요구조건에서는 빠른 페이지(FA블록 혹은 RA블록 내의 페이지)와 느린 페이지(SA블록 내의 페이지)를 번갈아 사용하여 빠른 페이지들의 낭비를 막는다.

FA블록과 RA블록은 페이지 사용 시점이 겹치게 되는데, 두 블록 중 우선도는 RA블록이 갖게 된다. 즉, 높은 성능이 요구될 때 RA블록이 존재한다면 우선적으로 RA블록 내의 MSB 페이지만이 사용되며, RA블록이 존재하지 않을 때만 FA블록 내의 LSB 페이지가 사용된다. 이는 FA블록은 모든 LSB 페이지를 사용한 후에도 SA블록으로 변환되어 MSB 페이지를 모두 사용한 이후에야 다시 fast 블록으로 사용될 수 있고, RA블록으로 변환될 수 있다 하더라도 시간적 지역성이 극도로 높은 경우를 제외하면 빠른 페이지를 다시 사용할 수 있는 블록이 되는 데 오랜 시간이 필요하기 때문이다. 즉, RA블록의 MSB 페이지를 우선적으로 처리하여 garbage collection의 대상이 되게 하고, 이를 통해 빠른 페이지들을 보다 원활하게 공급하기 위함이다.

3.4 통합에 따른 오버헤드

제안하는 기법은 기존 RPS 기법에서 제안했던 FTL 구조를 기반으로 하며[2], RBQueue의 관리만이 추가된다. SBQueue와 마찬가지로 블록 구조체의 리스트로써 이를 관리할 수 있으므로, 하나의 칩마다 해당 구조체를 관리한다는 가정 하에 칩 당 수 십 KB 규모의 메모리로 구현 및 유지가 가능하다. 또한, 페이지 할당 시 RBQueue의 탐색 및 블록 전환 과정이 추가되나, 수 백 us의 쓰기 명령 수행시간에 비해 해당 소프트웨어적 오버헤드는 거의 무시할 수 있는 수준이다. 결과적으로, 기존 RPS 기법에 비해 통합으로 인한 오버헤드는 극히 미미하다고 볼 수 있다.

4. 실험 결과

4.1 실험 환경

제안하는 기법의 유효성을 평가하기 위해, 3절에서 설명한 바 있는 통합 기법이 적용된 integrated-FTL을 시뮬레이션 환경에 구현하였다. 비교 FTL로서, MLC 낸드 플래시 메모리 기반의 일반적인 페이지 레벨 주소 사상(address mapping)을 지원하는 page-FTL과, page-FTL을 기반으로 SLP 기법과 RPS 기법이 각기 적용된

SLP-FTL과 RPS-FTL을 구현하였다. SLP-FTL은 쓰기 명령 인가 시, 모든 active 블록(active 블록은 칩 당 하나씩 할당) 중 LSB 페이지를 사용할 수 있는 블록을 찾고, 만약 가용한 LSB 페이지가 없을 경우, 대응하는 LSB 페이지가 무효화된 MSB 페이지를 찾아 SLP 기법의 적용을 도모한다. RPS-FTL은 integrated-FTL과 같이 칩 별로 FA 블록과 SA 블록을 할당하고, 높은 쓰기 성능이 요구될 때는 FA블록을, 반대의 경우에는 SA블록을 사용한다. 시뮬레이터 상 저장장치의 용량은 16GB이며, 네 개의 1GB 칩을 가진 네 개의 채널로 구성된다.

또한, 작업 부하의 특성에 따른 성능 개선 정도를 평가하기 위해, FIO 벤치마크 도구[5]를 사용하여 데이터 업데이트 비율을 달리한 250 MB의 쓰기 명령을 유티시간 없이 지속적으로 인가하였다. 예를 들어, 업데이트 비율이 100%일 경우, 250 MB크기의 한 파일을 반복하여 기록하는 작업 부하이며, 50% 업데이트 비율의 경우, 새로운 250 MB의 쓰기 중 50%는 기존에 저장된 파일에 대한 쓰기가 이뤄진다.

4.2 기법 별 쓰기 성능 개선 정도 비교

그림 6은 작업 부하의 업데이트 비율에 따른 각 FTL들의 쓰기 성능을 비교하여 보여준다. 그림 6에서 확인할 수 있듯, integrated-FTL은 대부분의 작업 부하에 대해 나머지 FTL보다 월등한 대역폭을 보인다(RPS-FTL 대비 평균 39%의 대역폭 향상). 특히, 업데이트 비율이 높은 경우, SLC 낸드 플래시 기반 저장장치와 유사한 성능 수준을 보이는데, 이는 인가되는 쓰기 명령을 높은 확률로 빠른 페이지 기록 방식을 적용하여 처리하기 때문이다.

보다 자세한 실험 결과의 분석을 위해, 그림 5와 같이 인가된 전체 쓰기 명령 중 빠른 페이지 기록 방식을 활용하여 처리된 쓰기 명령의 비율을 측정하였다. 그림 5에서 확인할 수 있듯, 업데이트 비율이 높을수록 integrated-FTL은 더욱 많은 쓰기 명령을 빠른 페이지 기록 방식

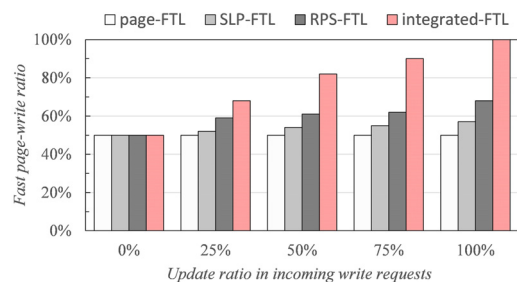


그림 5 전체 쓰기 명령 중 빠른 페이지 기록 방식이 적용된 명령의 비율 비교

Fig. 5 Comparisons of fast page-write ratio in incoming write requests for five I/O workloads

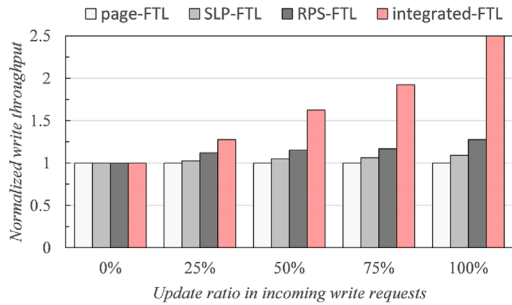


그림 6 FTL 별 쓰기 대역폭 비교

Fig. 6 Comparisons of normalized write throughput for five I/O workloads

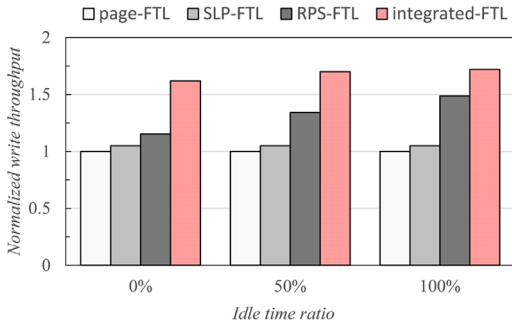


그림 7 유휴시간 변화에 따른 쓰기 대역폭 변화

Fig. 7 Comparisons of normalized write throughput for different idle time ratios

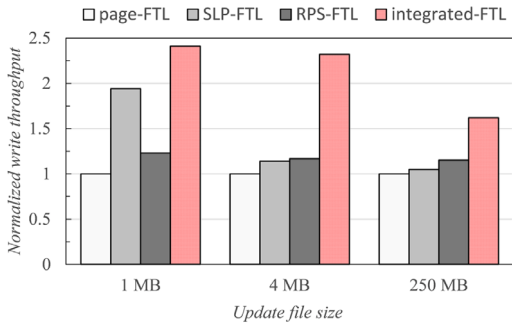


그림 8 업데이트 파일 크기 변화에 따른 쓰기 대역폭 변화
Fig. 8 Comparisons of normalized write throughput for different update file sizes

으로 처리하였다. 반면, 높은 업데이트 비율에도 불구하고 SLP-FTL과 RPS-FTL은 성능 개선이 크지 않은데, 이는 RPS-FTL의 경우 쓰기 명령이 지속적으로 인가되기 때문에 결국 LSB 페이지와 MSB 페이지를 번갈아 사용하게 되고, SLP-FTL의 경우, LSB 페이지가 무효화될 확률이 매우 낮기 때문에 기법의 적용 가능성이

매우 떨어지기 때문이다.

두 기법의 통합에 의한 효과를 보다 자세히 확인하기 위한 실험으로, 50%의 업데이트 비율을 갖는 작업부하에 대해 유휴시간과 업데이트되는 파일 크기를 조정해가며 실험을 진행하였다. 유휴시간은 250MB의 쓰기 명령이 진행될 때마다 수행 시간의 50%, 100%의 시간 동안 쓰기명령을 인가하지 않음으로써 조정하였고, 업데이트 파일의 크기는 1MB, 4MB의 경우를 추가하여 실험하였다. 그림 7에서 확인할 수 있듯, RPS-FTL은 유휴시간이 많은 경우 백그라운드 가비지 컬렉션에서 느린 MSB페이지를 사용할 수 있어 쓰기 명령이 집중되는 동안 더 많은 LSB 페이지를 사용함으로써 성능을 개선할 수 있다. 반면, 유휴 시간이 없을 때는 LSB 페이지와 MSB 페이지를 번갈아 사용하므로 성능 개선이 대폭 감소함을 확인할 수 있다. 한편, 그림 8에서 확인할 수 있듯, SLP-FTL은 1MB의 범위 내에서 지속적인 업데이트가 발생할 경우, 유휴시간이 없더라도 큰 성능 향상을 보일 수 있지만, 범위가 조금만 확장되더라도 LSB페이지의 무효화 확률이 크게 떨어짐으로 인해 성능 개선이 거의 없음을 확인할 수 있다. 반면, integrated-FTL에서는 업데이트의 범위가 넓고 유휴시간이 없더라도, 업데이트 비율이 일정 수준 존재한다면 상당한 성능 개선이 가능하다는 것을 확인할 수 있다.

5. 결론

다치화 기법은 낸드 플래시 메모리의 집적도 향상에 지대한 공헌을 했지만, 그로 인해 기존 SLC 낸드 플래시 메모리보다 악화된 MLC 낸드 플래시 메모리의 성능은 저장장치에 대한 높은 성능 요구조건을 만족하기 위해 필수적으로 해결해야 할 문제로 대두되었다. 본 논문에서는 MLC 낸드 플래시 기반 저장장치의 쓰기 성능을 개선하기 위한 기존 계층 교차적 최적화 기법들의 한계점을 분석하고, 해당 기법들의 상호 보완적인 측면을 이용하여 효율적인 통합을 통해 그 한계점을 극복하는 새로운 기법을 제안하였다. 실험 결과, 제안하는 기법은 MLC 낸드 플래시 기반 저장장치의 쓰기 성능을 평균 39% 개선할 수 있음을 확인하였다.

References

- [1] C. Kim et al., "A 21 nm High Performance 64 Gb MLC NAND Flash Memory with 400 MB/s Asynchronous Toggle DDR Interface," *Proc. IEEE International Solid-Sate Circuits Conference*, 2013.
- [2] J. Park et al., "Improving Performance and Lifetime of NAND Storage Systems Using Relaxed Program Sequence," *Proc. Design Automation Conference*, 2016.

- [3] Y.-M. Chang et al., "Achieving SLC Performance with MLC Flash Memory," *Proc. Design Automation Conference*, 2015.
- [4] K.-T. Park et al., "A Zeroing Cell-to-Cell Interference Page Architecture with Temporary LSB Storing and Parallel MSB Program Scheme for MLC NAND Flash Memories," *IEEE Journal of Solid-State Circuits*, Vol. 43, No. 4, pp. 919-928, Apr. 2008.
- [5] Fio benchmark. [Online]. Available: <http://freecode.com/projects/fio> (accessed 2017, Feb. 10)



박 지 성

2011년 서울대학교 컴퓨터공학부 학사
2011년~현재 서울대학교 석박통합과정
관심분야는 임베디드 시스템, 컴퓨터 구조,
플래시 저장장치



이 성 진

2011년 고려대학교 전자공학 학사. 2007
년 서울대학교 컴퓨터공학부 석사. 2013년
서울대학교 컴퓨터공학 박사. 2013년~
2016년 MIT CSAIL 박사후연구원. 2016
년~2017년 인하대 컴퓨터 공학과 조교
수. 2017년~현재 DGIST 정보통신융합
전공 조교수. 관심분야는 운영체제, 컴퓨터 구조, 임베디드
시스템, 플래시 저장장치

김 지 홍

정보과학회논문지
제 44 권 제 3 호 참조