낸드 플래시 메모리 시스템의 Lifetime 증대를 위한 관리 방법 설계

박이현*, 이재빈*, 김건명*, 임승호*
*한국외국어대학교 컴퓨터.전자시스템공학부
e-mail:slim@hufs.ac.kr

NAND Flash Memory System Management for Lifetime Extension

Yi-Hyun Park*, Jae-Bin Lee*, Geon-Myung Kim*, Seung-Ho Lim*
*Division of Computer and Electronic Systems Engineering
Hankuk University of Foreign Studies

요 약

낸드 플래시 메모리(NAND Flash Memory)는 컴퓨터 시스템의 대용량 저장장치를 위한 소자로써, 대용량화의 주요 원인으로는 메모리 셀(Cell) 당 저장할 수 있는 비트 수를 증가시킴으로써 집적도를 증가시킨 것이다. 그러나, 이러한 집적도의 증가는 에러의 증가를 가져와서 저장장치에서 가장 중요한 신뢰성이 급격하게 저하하는 요인이며, 저장장치의 생명주기(Lifetime)을 감소시키게 된다. 기존에 낸드 플래시 메모리 저장장치의 Lifetime을 증대시키기 위해서 P/E cycle을 고려하여 데이터 영역의 일부를 점점 더 ECC 영역으로 변경시키는 방식을 적용한 바가 있다. 이러한 방식은 데이터 영역의 감소로 인한 저장장치 내에서 관리되는 호스트-플래시 간 데이터 관리 크기의 미스매치로 인한 여러 가지 오버헤드를 생성한다. 본 연구에서는 P/E cycle에 따른 데이터 영역의 ECC 영역으로의 전환을 통한 Lifetime을 증가시키는 방식에 있어서, 오버헤드를 줄이기 위한 캐쉬 관리 구조 및 매핑 관리 구조에 대한 설계를 진행하였다. 이러한 설계를 낸드 플래시 메모리 기반 저장장치에 적용할 경우, LifeTime을 증대시키기 위해서 ECC를 데이터 영역으로 확장하는 방식을 사용할 때 저하될 수 있는 일반 읽기 및 쓰기의 성능 저하를 어느 정도 감소시켜줄 수 있을 것으로 기대한다.

1. 서론

낸드 플래시 메모리(NAND Flash Memory)는 컴퓨터 시스템의 대용량 저장장치를 위한 소자로써, 낸드 플래시 메모리 기반의 SSD 및 eMMC 등은 컴퓨터 시스템에서 현재 널리 사용되고 있으며 그 활용성은 점점 더 증폭될 것으로 예상된다. 낸드 플래시 메모리의 대용량화의 주요 원인으로는 메모리 셀(Cell) 당 저장할 수 있는 비트 수를 증가시킴으로써 집적도를 증가시키고 있다. 현재 널리 사용되는 메모리 셀은 TLC(Three Level Cell)로써 셀 당 세 비트를 저장하며, QLC(Quad Level Cell) 등 셀 당 비트 수는 점점 증가하는 추세이다. 그러나, 이러한 집적도의 증가로 인해 셀 당 비트 간의 간섭을 증가시켜 에러를 발생시키는 확률이 급격히 증가하고 있으며, 집적도의 증가는 에러의 증가를 가져와서 저장장치에서 가장 중요한 신뢰성을 급격하게 저하시키는 요인이며, 저장장치의 생명주기(Lifetime)을 감소시키게 된다[2.3].

낸드 플래시 메모리의 에러를 보정하고, 저장장치의 에러율을 감소시키기 위해서, 낸드 플래시 메모리는 BCH 또는 LDPC와 같은 ECC(Error Correction Code)[1,4]를 이용하여, 데이터 영역에 대한 부가적인 ECC 데이터를 생성하여 별도로 저장하고 있다. 그러나, 낸드 플래시 메모리에서 발생하는 에러는 낸드 플래시 메모리의 P/E cycle이 증가할수록 에러가 발생하는 확률이 점점 더 증대하게 되어, 고정된 데이터 영역에 대한 고정된 ECC만으로는 P/E cycle 대비 저장장치의 Lifetime이 짧아지는 결과가 나타난다.

기존의 연구에서는 낸드 플래시 메모리 저장장치의 Lifetime을 증대시키기 위해서 P/E cycle을 고려하여 데이터 영역의 일부를 점점 더 ECC 영역으로 변경시키는 방식을 적용한 바가 있다[5]. 이러한 방식은 ECC 영역의 증대로 인해서 Lifetime은 증가하지만, 데이터 영역의 감소로 인한 저장장치 내에서 관리되는 호스트-플래시 간 데이터 관리 크기의 미스매치로 인한 여러 가지 오버헤드를 생성한다. 본 연구에서는 P/E cycle에 따른 데이터 영역의 ECC 영역으로의 전환을 통한 Lifetime을 증가시키는 방식에 있어서, 오버헤드를 줄이기 위한 캐쉬 관리 구조 및 매핑 관리 구조에 대한 설계를 진행하였다.

2. Lifetime 증대를 위한 플래시 관리 기법

낸드 플래시 메모리는 읽기/쓰기/지우기 연산을 통해서 사용된다. 읽기/쓰기는 페이지(Page)라는 연산단위로 동작하며, 지우기는 블록(Block)이라는 연산단위로 동작한

다. 일반적으로 수십 개의 페이지(예, 128개)는 하나의 블록을 구성한다. 페이지 내부는 또다시 호스트의 데이터를 저장하는 데이터 영역(Data Region)과 ECC와 같은 데이터의 부가적인 정보를 저장하기 위한 스페어 영역(Spare Region)영역으로 나뉘며, 이 크기는 낸드 플래시 메모리 제조 공정 시에 고정되어 있다. 일반적으로 ECC 데이터가 스페어 영역에 저장되므로 스페어 영역을 ECC 영역으로 봐도 무방하다. 낸드 플래시 메모리는 쓰기가 발생하기 전에 지우기 연산을 수행해 주어야 하며, 이 한 번의 주기를 P/E cycle이라 한다. 낸드 플래시 메모리 초기 생명 주기인 P/E cycle 이 낮을 경우 메모리 셀에서 에러가 발생할 확률이 적으며, P/E cycle이 늘어날수록 메모리 셀에서 에러가 발생할 확률이 높아진다. 따라서, 초창기 P/E cycle의 적은 ECC로도 에러 발생 시 복원이 가능하지만, 점점 P/E cycle이 증가할수록 에러 복원에 필요한 ECC 영역은 점점 더 많이 필요하게 된다.

본 연구에서는 P/E cycle이 증가할수록 에러 복원확률을 높이기 위해서 페이지의 데이터 영역을 점점 줄이고, ECC 영역을 점점 증대하는 방식의 낸드 플래시 생명 주기 관리 방법에 대한 시스템 설계를 진행하였다. 컴퓨터 시스템에서 호스트와 저장장치 간의 데이터 읽기/쓰기는 일반적으로 정의된 크기 단위 혹은 크기 단위의 배수로 읽기/쓰기 연산이 발생하며, 이 크기는 호스트-저장장치 간의 연산뿐만 아니라 저장장치 내부에서의 연산의 효율성을 위해서 낸드 플래시 메모리의 페이지 크기와 같거나 정수배로 고정된다. 그런데, P/E cycle이 증가함에 따라서 ECC 영역의 크기를 늘리기 위해서 데이터 영역을 줄이게 되면, 호스트에서 전송하는 데이터가 페이지에 딱 맞게 저장되지 않은 분할(Fragmentation)이 발생한다. 기존의 연구에서는 이러한 분할이 발생하는 경우 단순히 다음 페이지를 할당하여 연속적으로 쓰는 방식을 적용하였는데, 이 경우 읽기/쓰기에서 매번 분할로 인한 부하가 발생하게 된다.

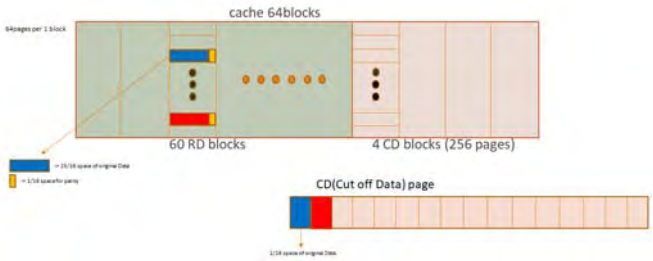

그림 1 ECC 확장에 의해서 분할된 데이터 관리 캐쉬 구조

이러한 분할로 인한 시스템 부하를 줄여주기 위해서, 본 연구에서는 데이터 분할로 인해서 남는 데이터 조각들에 대해서 따로 영역을 할당하여 캐싱하고 저장하는 방식에 대한 설계를 진행하였다. 본 설계에서는 P/E cycle이 증가함에 따라서 단계적으로 데이터 영역을 줄이고, 줄어든 데이터 영역 만큼 분할 된 데이터를 따로 모아, 여러 개의 페이지의 분할된 부분을 하나의 페이지로 만들어 캐쉬에 관리한다. 그림 1은 데이터 분할이 발생한 영역(Reduced Data, RD)과 분할 된 데이터를 모아서 따로 저장하는 영역(Cut-off Data, CD)영역의 캐쉬 구조도를 예를 들어 도식화 한 것이다. 예를 들어, ECC를 위해서 추가적으로 할당한 데이터 영역의 크기가 원래 데이터 영역의 1/16이라고 할 경우, 하나의 RD block에 대해서 연속된 4개의 CD pages 가 배정된다. 4개의 CD pages의 가장 첫 page의 index 번호가 특정 RD block에 부여되어 해당 RD block에 write가 발생하면 배정된 CD pages에 Cuf off Data가 write된다. 만약, 다른 Block index 를 가지는 RD block에 write가 일어난다면 새로운 4개의 CD pages를 할당하여 CD를 write 하도록 한다.

RD block 하나에 배정된 4개의 CD pages는 한 page 당 16개, 총 64개로 분할되는 영역이 제공되는데, 64개로 분할된 CD page 영역은 하나의 RD block 내부 64개 pages에 저장된 data 들과 각각 동일한 lpn 이 저장된다.

분할된 데이터 영역으로 인해서 발생하는 쓰기 오버헤드를 줄이기 위해서 CD가 저장된 블록의 리스트와 RD가 저장된 블록의 리스트를 따로 관리해주고, 캐쉬에서 replace가 발생하더라도 별도로 victim 선정이 적용되도록 함으로써 CD 데이터의 쓰기 지연과 캐싱이 가능하도록 하였다. 즉, 호스트의 쓰기 발생으로 인해서 더이상 캐쉬에 저장할 공간이 없는 경우, 캐쉬의 replace가 발생한다. 기본적으로 replace의 victim 선정은 RD block들 사이에서 선정되게 되며, RD block replace 시 선택된 victim block은 replace 하더라도, 해당 블록과 연관된, 즉 분할된 데이터가 캐싱된 해당 CD pages의 data는 flush 되지 않는다. CD block의 replace는 다음과 같다. 만약, CD block의 256 pages가 꽉 차게 되면, 자체 CD block list의 LRU 방식에 의하여 4개의 연속된 CD pages를 chip 내부 메모리에 저장시켜주고 저장시킨 CD pages와 연관된 RD block의 페이지들을 chip 내부 메모리에 저장시킨 후, cache에서 연속된 4개의 CD pages를 제거함으로써 CD를 위한 새로운 영역을 확보한다.

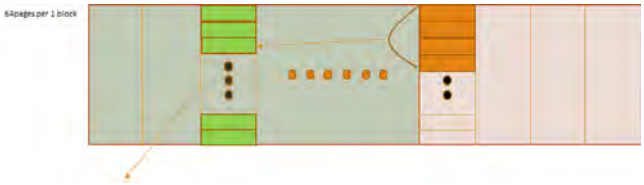

그림 2 분할된 데이터 영역(Reduced Data)과 분할 데이터(Cut-off Data)와의 연계 및 상관관계

DRAM 캐쉬에 분리되어 캐싱된 데이터는 캐쉬 replace 정책에 의해서 플래시 메모리 쓰기 연산으로 변환되어 플래시 메모리에 프로그래밍 되며, 플래시 메모리 영역에 저장되는 방식과 이를 위한 FTL 매핑 관리 구조는 다음과 같다. 먼저, Flash memory block 내부 block들은 해당 block이 RD block 인지 CD block 인지를 나타내는 block_state라는 field가 추가 부여된다. 그리고, 새로운 RD block에 대한 쓰기 시, flash memory 상에 저장되는 16개의 연속된 RD page들은 RD page의 ppn 에 따라 CD block 내의 하나의 CD page의 index를 가리키도록 FTL에 매핑 정보를 기록한다. flash memory read 시 특정 ppn 의 data를 읽는다면, CD mapping table에 의하여 CD page의 index를 확인하여 해당 CD page의 data를 함께 읽어 들인다. 그림 3은 플래시 메모리 내에서 RD block과 CD블록 사이의 연관성을 도식화 한 것이며, 이러한 정보는 FTL 매핑 테이블에서 관리된다. 예를 들어 RD block의 3번째 page(2번 데이터)를 read 시, 해당 CD page의 index 2 에 해당하는 data를 read 한다. 이를 위해 CD page의 index는 page 당 0~15가 부여된다. flash memory write 시에도 연결된 CD page의 정보를 기반으로 해당 CD page에 index에 따라 write 한다.

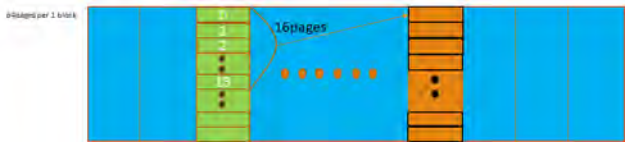

그림 3 플래시 메모리 내의 RD block과 CD block 저장 구조도

3. 결론

낸드 플래시 메모리 기반의 저장장치는 셀의 집적도가 높아짐에 따라서 셀 당 에러가 발생하는 확률도 점점 증가하여, 저장장치의 신뢰성에 대한 문제가 중요한 이슈가 되고 있다. 낸드 플래시 메모리 기반의 저장장치의 수명을 증대시키기 위해서 P/E cycle 변화에 따른 ECC 양을 동적으로 증가시켜 에러 복원율을 높이는 방식을 적용할 수 있다. 그러나 이러한 방식은 기존의 데이터 영역에 저장되는 데이터의 파편화를 가져올 수 있으며 이는 연산의 오버헤드를 발생시킨다. 본 논문에서는 데이터 영역의 파편화에 따른 연산 오버헤드를 감소시켜줄 수 있는 파편화 데이터 영역의 캐싱 관리 방법과 플래시 메모리 저장 관리 방법에 대한 설계를 진행하였다. 이러한 설계를 적용함으로써 P/C cycle 대비 ECC 영역의 증가로 인한 데이터 연산의 파편화 오버헤드를 감소시킬 수 있으며, 저장장치의 수명 연장과 동시에 성능 감소를 최소화시킬 수 있을 것으로 기대한다.

Acknowledgments

The Research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (NRF-2016R1C1B2009914).

참고문헌

[1] S. Lin and D. J. Costello, “Error Control Coding: Fundamentals and Applications,” Englewood Cliffs, New Jersey, 1983.

[2] Intel, “JS29F16B08CCME2 MLC NAND Flash Memory Datat Sheet,” 2011.

[3] Y. Cai, E. H. Onur.Mutlu, and K. Mai, “Error patterns in MLC NAND flash memory: Measurement, characterization, and analysis,” in Proc. of Design, Automation & Test in Europe Conference & Exhibition (DATE), pp. 521 - 526, 2012.

[4] K. Zhao, W. Zhao, H. Sun, X. Zhang, N. Zheng, and T. Zhang, “LDPCin-SSD: Making advanced error correction codes work effectively in solid state drives,” in Presented as part of the 11th USENIX Conference on File and Storage Technologies (FAST), pp. 243 - 256, 2013.

[5] Shunzhuo Wang. et al. “Lifetime adaptive ECC in NAND flash page management”, in Proceedings of Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017