

SSR: 3D NAND 플래시 메모리 기반 대용량 저장장치에 특화된 중단/재개 기법

김 범 준*, 서 경 섭*, 김 명 석^o

SSR: Smart Suspend/Resume Technique for High Capacity 3D NAND Flash Memory Based Storage System

Beomjun Kim*, Gyeongseob Seo*, Myungsuk Kim^o

요 약

NAND 플래시 메모리 기술이 지속적으로 확장됨에 따라 플래시 기반 SSD는 최신 스토리지 시스템의 핵심 구성 요소가 되었다. 최신 스토리지 시스템의 주요 설계 목표 중 하나는 데이터 센터, 클라우드, LLM 및 자율자동차와 같은 최첨단 애플리케이션에 필수적인 낮은 읽기 응답시간을 달성하는 것이다. 이러한 목표를 달성하기 위해 이전의 많은 연구는 읽기 작업에 대한 요청이 있을 때 현재 진행 중인 지우기 작업을 즉시 중단하고 읽기 요청을 처리함으로써 읽기 응답시간을 최소화하였다. 안타깝게도 이러한 지우기 중단 기술은 임의의 지점에서 지우기 동작을 중단/재개하기 위해 NAND 주변 장치에 추가 하드웨어 비용을 발생시킬 뿐만 아니라 무엇보다 SSD의 수명을 심각하게 단축시킨다. 또한, 기존의 중단/재개 기술은 쓰기 동작의 기아 문제(starvation)를 피할 수 없다. 이러한 제한을 극복하기 위해 본 논문에서는 최신 SSD에서 사용되는 반복적 지우기 메커니즘을 활용하는 새로운 시스템 수준 지우기 중단 기술인 SSR(Smart Suspend Resume)을 제안한다. SSR은 160개의 실제 3D TLC NAND 플래시 칩을 사용하여 구성된 정밀한 플래시 메모리 특성 모델링을 기반으로 단일 erase loop 내에 여러 개의 safe point를 도입하고, 읽기 요청 인가 시 safe point에서만 중단/재개 동작을 지원하도록 한다. 이를 통해 중단/재개 동작의 발생 빈도를 제한함으로써 저장장치의 신뢰성 저하와 쓰기 기아 문제를 효과적으로 방지할 수 있다. 시스템 수준의 다양한 평가 결과, SSR은 기존 기법보다 읽기 응답시간의 지연을 최대 10배 개선할 수 있음을 보여준다.

키워드 : 낸드 플래시 메모리, 저장장치, SSD, 중단, 재개, 지우기 동작

Key Words : NAND flash memory, Storage system, SSD, Suspend, Resume, Erase

ABSTRACT

As NAND flash technology continues to scale, flash-based SSDs have become key components in storage systems. One of the main design goals for modern storage systems is low read tail latency, which is crucial for state-of-the-art applications such as data centers or LLM. Towards this goal, prior works on erase suspension address this problem by allowing a read operation to interrupt an ongoing erase operation.

* 본 연구는 과학기술정보통신부 (RS-202400414964), 교육부 (NRF-2022R1I1A3073170) 및 정보통신기획평가원 (RS-2024-00347394)의 지원에 의해 수행되었습니다.

• First Author : Kyungpook National University, School of Computer Science and Engineering, beomjun0816@knu.ac.kr, 학생회원; syhbong9@knu.ac.kr, 학생회원

* Corresponding Author : Kyungpook National University, School of Computer Science and Engineering, ms.kim@knu.ac.kr, 정회원
논문번호 : 202501-006-E-RU, Received January 2, 2025; Revised February 25, 2025; Accepted February 28, 2025

Unfortunately, the conventional technique attempts to suspend/resume an erase pulse at an arbitrary point, which incurs additional hardware costs for NAND peripherals and significantly reduces the SSD's lifetime. Furthermore, the previous technique suffers from an erase/write starvation problem. To overcome these limitations, we propose a new suspend/resume technique, SSR (Smart Suspend Resume), leveraging the ISPE mechanism used in modern SSDs. The SSR introduces multiple safe points within a single erase loop based on a device model using 160 real 3D TLC flash chips, and performs suspend/resume erase operations at well-aligned safe points. By doing so, we can limit the frequency of suspend/resume operations, thus effectively avoiding reliability degradation and erase/write starvation problems. Comprehensive evaluations reveal that our SSR can improve the read tail latency up to 10 times over the baseline for two popular workloads.

I. 서 론

NAND 플래시 기반 SSD는 하드 디스크에 비해 뛰어난 처리량과 낮은 평균 응답 시간(latency)을 제공하는 장점으로 인해 최신 저장장치의 실질적인 표준이 되었다. 그러나 SSD는 하드 디스크와 비교하면 성능의 변동성이 크다는 문제를 안고 있다^[1]. 예를 들어, SSD는 매우 낮은 평균 읽기 응답시간(예를 들어, 15 μ s 미만^[2])을 달성할 수 있지만, 다양한 원인으로 인해 매우 긴 꼬리 응답시간(tail latency)을 보이기도 한다. 그림 1은 실제 SSD에서의 응답시간 지연 문제를 보여준다. 예를 들어 전체 읽기 요청의 90%에 대해서는 최소 20 μ s에서 최대 120 μ s의 응답속도의 분포를 보인다. 예시의 PCIe 3 \times 4 NVMe SSD는 읽기 요청에 대하여 100 μ s 평균 응답시간을 갖지만, 전체 읽기 요청의 5%는 400 μ s(95% 백분위), 1%는 800 μ s(99% 백분위), 0.01%는 9000 μ s(99.99% 백분위), 그리고 99.999% 백분위에 해당하는 0.001%의 읽기 요청은 10ms를 초과하는(평균 응답시간의 1000배 수준) 응답시간을 보인다. 이러한 응답시간을 평균 응답시간과 대비하여 꼬리 응답시간이라고 하며 긴 꼬리 응답시간은 호스트의 읽기 요청에 대한 저장장치의 응답 지연이라는 심각한 문제를 유발한다.

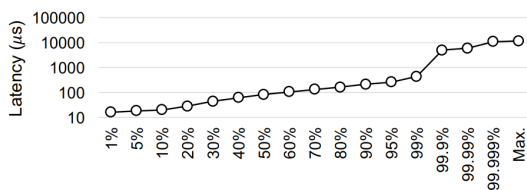


그림 1. 대기열 깊이(queue depth) 16인 4KB 무작위 읽기(70%) 및 쓰기(30%) 워크로드를 실행하는 상용 PCIe 3 \times 4 NVMe SSD의 읽기 요청에 대한 응답속도(latency) 분포
Fig. 1. Read latency distribution of a commercial PCIe 3 \times 4 NVMe SSD running 4KB random reads (70%) and writes (30%) workload with a queue depth of 16.

이러한 읽기 작업에 대한 긴 꼬리 응답시간의 문제는 최신 데이터 중심 컴퓨팅환경에서 저장장치뿐만 아니라 컴퓨팅 시스템 전체의 성능을 좌우하는 중요한 평가 기준이 되고 있다. 최신 데이터센터의 경우 단일 query (예를 들어, 웹 검색)가 데이터센터를 구성하는 수천 개의 디스크에 대한 읽기 작업을 요청할 수 있으며 이러한 경우 꼬리 응답시간으로 인한 지연 문제는 전체 query 응답시간의 심각한 증가를 유발한다. 특히 생성형 AI, 유전체 분석, 및 자율 운행 자동차와 같이 대규모의 데이터를 사용하는 최신 응용에서^[3,4] 긴 꼬리 응답시간이 시스템의 성능에 끼치는 영향은 더욱 증대하고 있다. 이러한 경향을 반영하여 산업계에서는 저장장치의 성능을 4-nine (99.99%) 또는 6-nine (99.9999%) 백분위의 응답시간을 활용한 QoS 지표를 중요한 성능지표로 삼고 있다^[5].

이러한 긴 꼬리 응답시간의 문제를 해결하기 위해서는 응답시간의 지연을 유발하는 원인을 제거하는 것이 필요하다. NAND 플래시 메모리 기반 SSD에서 응답시간의 지연은 주로 읽기 요청이 상대적으로 긴 응답시간을 갖는 쓰기(1ms-2ms) 또는 지우기(3.5ms-20ms) 동작과 충돌하여 읽기 작업이 지연됨으로써 발생한다. 읽기 요청이 쓰기(program) 또는 지우기(erase) 동작 때문에 지연되는 현상은 사용자가 요청하는 읽기/쓰기/지우기 동작 사이의 충돌에 의해 발생하기도 하지만 주로 가비지 컬렉션(Garbage Collection, GC), 웨어-레벨링(Wear Leveling)과 같은 저장장치 내부를 관리하기 위한 작업과 호스트에서 인가하는 사용자 읽기 작업과의 충돌로 인해 발생한다. 특히 빈번하게 발생하는 GC는 SSD 읽기 응답시간의 지연을 유발하는 주요 원인 중 하나이다^[6,7]. NAND 플래시 메모리는 HDD와는 다르게 데이터를 덮어쓸 수 없으므로 데이터 쓰기 동작을 수행하기 위해서는 모든 데이터가 지워진 빈 블록(free block)이 확보되어야 한다¹⁾. 만약 free block이 없으면 기존 블록들에서 유효한 데이터(valid data)만을 복사하

여 하나의 블록에 모으고 기존 블록들에 대해서 지우기 동작을 수행하여 free block을 생성하게 되며 이러한 작업을 GC라고 한다. GC는 유효한 데이터에 대한 읽기와 쓰기 동작, 그리고 free block들을 만들기 위한 지우기 동작으로 구성되어 있으므로 굉장히 긴 수행 시간(예를 들어 100ms 이상)이 필요하며 읽기 작업이 GC와 충돌하게 되면 GC가 완료될 때까지 읽기 작업의 응답 시간은 지연될 수밖에 없다. 따라서 꼬리 응답시간의 문제를 해결하기 위해서는 사용자의 읽기 요청과 쓰기 또는 지우기 동작이 충돌할 시 쓰기 또는 지우기 동작을 일시적으로 멈추고 사용자의 읽기 작업을 우선 수행하는 플래시 메모리 기반 중단/재개(suspend/resume) 동작의 구현이 필수적이며 지난 십여 년 동안 쓰기과 지우기 동작에 대한 다양한 중단/재개 기법이 연구되었다^[8-11].

쓰기 동작에 대한 중단/재개 기법은 NAND 플래시 메모리 쓰기 동작에서 활용되는 ISPP(Incremental Step Pulse Programming)를 활용한 다양한 기법이 제안되었다^[8,9]. NAND 플래시 메모리는 쓰기 동작 시 데이터의 비트 에러를 최소화하기 위해 한 번에 높은 전압으로 쓰기 동작을 하지 않고 전압을 조금씩 높여가면서 여러 번에 나누어(최종 쓰기 완료까지 20회 이상) 쓰기 동작을 수행하며 개별 동작을 program loop라고 한다. 기존의 기법들은 쓰기 동작에 대한 중단/재개 기법을 개별 program loop 단위로 수행함으로써 신뢰성 열화 없이 꼬리 응답시간을 개선하였다. 즉, 쓰기 동작 수행 중에 읽기 요청이 발생하는 경우 해당 읽기 요청을 바로 수행하는 것이 아니라 하나의 program loop가 완료된 후 읽기 요청을 처리함으로써 읽기 동작의 지연 시간을 program 1 loop 수행 시간(약 100 μ s) 이내로 줄이면서도 쓰기 동작이 무한정 지연되지 않고 일정 시간 이내에는 완료되는 것을 보장해준다.

지우기 동작의 경우 두 가지 이유로 쓰기 동작보다 훨씬 더 정교한 중단/재개 기법이 필요하다. 첫 번째, 지우기 동작의 응답시간은 쓰기 동작보다 수 배에서 수십 배 크다^[12]. 따라서 읽기 요청이 지우기 동작과 충돌하는 경우 더 치명적인 꼬리 응답시간의 열화가 발생한다. 두 번째, 지우기 동작은 NAND 플래시 메모리의 수명을 결정짓는 핵심 요소이다^[13]. 따라서 지우기 동작에 대한 중단/재개 기법으로 인해 불필요한 스트레스가 NAND 플래시 메모리에 인가되면 SSD와 같은 저장장치의 수명은 급격히 줄어들게 된다.

지우기 동작에 대한 중단/재개 기법으로 Wu et al.는 지우기 동작이 tail latency에 미치는 영향을 제거하기 위해, 읽기 요청이 발생하면 그 즉시 진행 중인 지우

기 동작을 중단하고 읽기 요청을 우선 처리하고 그 이후 중단되었던 시점 이후부터 지우기 동작을 재개하는 I-ES (immediate Erase Suspension) 기법을 제안하였다^[10]. I-ES는 읽기 요청의 지연을 제거하는 데 효과적이지만 세 가지의 심각한 문제를 야기한다. 첫 번째, 읽기 요청으로 인한 중단이 발생할 때마다 중단된 지점을 저장하고 재개 동작 이후 정확히 남은 지우기 동작을 수행하기 위해 추가적인 설계 비용이 발생한다. 두 번째, 지우기 동작 수행 중 빈번한 중단/재개 동작이 발생하는 경우 over-erase가 발생하여 NAND 플래시 메모리의 수명을 심각하게 훼손하게 된다. 마지막으로, 읽기 요청이 극단적으로 많이 발생하는 경우 오히려 지우기 동작의 수행이 지연되어 쓰기 동작의 기아(starvation) 현상이 발생할 수 있다²⁾. 이를 보완하기 위해서 Kim et. al.은 쓰기 동작의 중단/재개 기법과 유사하게 읽기 요청에 대한 처리를 개별 erase loop가 완료된 후 수행하는 D-ES (Deferred Erase Suspension) 기법을 제안하였다^[11]. D-ES는 쓰기 동작의 기아 문제를 해결하고 읽기 요청의 지연 시간을 일정 수준 이하로 개선할 수 있다. 하지만 지우기 동작의 1 erase loop 수행 시간이 수 ms에 달하기 때문에 D-ES가 읽기 꼬리 응답시간의 문제를 개선하는 데에는 한계가 있다.

본 논문에서 우리는 최신 3D NAND 플래시 메모리의 지우기 동작에 특화된 새로운 중단/재개 기법을 (SSR, Smart Suspend Resume) 제안하고자 한다. 우리는 160개의 실제 3D TLC NAND 플래시 칩을 활용하여 지우기 동작에 인가되는 중단/재개 동작으로 인한 신뢰성 열화를 모델링 하였다. 이러한 모델링을 바탕으로 신뢰성 열화 없이 수용 가능한 중단/재개 동작의 횟수를 확인하고 이를 활용하여 단일 erase loop 동작을 나누어 중단/재개 동작을 안전하게 지원할 수 있는 구간을 설정하였다. 이러한 접근을 통해 SSR은 신뢰성 열화와 쓰기 동작의 기아 현상을 회피하면서 SSD 읽기 동작의 꼬리 응답시간의 개선을 극대화하였다.

해당 기법의 효과를 검증하기 위해 가장 널리 사용되는 시스템 수준의 저장장치 시뮬레이터 중 하나인 MQSim^[14]을 활용하여 SSR을 지원하는 Smart-FTL을 구현하였다. 다양한 워크로드를 사용한 실험 결과 Smart-FTL은 기존 기법에 비해 99.99%와 99.9999%

1) NAND 플래시 메모리 칩은 물리적으로 구분된 수천 개의 블록으로 구성되어 있다^[12].

2) 쓰기 동작을 수행하려면 그 전에 해당 블록에 대한 지우기 동작이 선행되어야 한다. 이를 NAND 플래시 메모리의 고유한 erase-before-write 특성이라고 한다. 따라서 지우기 동작의 완료가 지연되면 쓰기 동작을 수행할 수 없게 된다.

기준 저장장치의 꼬리 응답시간을 신뢰성 열화 없이 최대 10배 개선하였다.

II. 배경 지식

2.1 NAND 플래시 메모리 개요

그림 2는 NAND 플래시 칩의 구조를 계층적으로 보여준다. 플래시 셀은 비트라인(BL)에 연결되어 NAND 스트링을 형성하고, 서로 다른 BL(8K개 또는 16K개)의 NAND 스트링이 모여 블록을 구성한다. 그림 2의 개별 플래시 블록에는 m 개의 WL이 있고 각 WL은 n 개의 플래시 셀(n 개의 BL)로 구성되어 있다. 블록의 동일한 수직 위치에 있는 개별 플래시 셀은 워드라인(WL) 게이트를 공유하기 때문에 동일한 WL에 있는 모든 셀은 읽기, 쓰기 동작을 동시에 수행하게 된다. 개별 BL은 데이터 입/출력 회로를 통한 데이터 전송을 담당하는 페이지 버퍼에 연결된다. 결론적으로 플래시 셀은 수평 방향으로 WL, 수직 방향으로 BL을 공유하며 매트릭스 구조를 형성한다.

플래시 셀은 구조적으로 일반적인 MOS 트랜지스터와 유사하지만, 내부 특정 영역에(트랩 층, trap layer) 전자를 주입 또는 제거하여 임계전압(threshold voltage, V_{th})을 조절할 수 있다는 점에서 차이가 있다. 주입된 전자의 수에 따라 플래시 셀은 서로 다른 V_{th} 값을 가지게 되며 특정 게이트 전압에서 off 또는 on 스위치로 작동하여 데이터를 저장한다(즉, 주입된 전자의 수가 많을수록 V_{th} 의 값이 증가하게 되며 비트 데이터를 V_{th} 값의 함수로 인코딩하여 셀에 저장한다). 예를 들어, 높은 V_{th} 값을 갖는 플래시 셀에 데이터 '0'을 할당하고, 낮은 V_{th} 값을 갖는 플래시 셀에 데이터 '1'을 할당함으로써 1 비트의 정보를 성공적으로 저장할 수 있다.

NAND 플래시 메모리는 저장된 데이터에 액세스하거나 수정하기 위해서 세 가지 기본 동작 (i) 쓰기, (ii) 읽기, (iii) 지우기 동작을 지원한다. 선택된 플래시 셀의

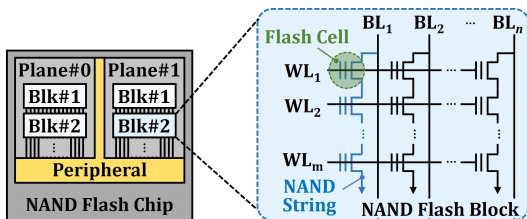


그림 2. NAND 플래시 메모리의 개략적인 구조
Fig. 2. An organizational overview of NAND flash memory.

V_{th} 를 증가시키는 쓰기 동작은 WL 게이트에 높은 전압($>20V$)을 인가함으로써 FN 터널링을^[15] 유발하여 기판으로부터 플래시 셀의 전하 트랩 층으로 전자를 주입한다(즉, V_{th} 를 높여 플래시 셀의 상태를 '1' 상태에서 '0' 상태로 변경). 반면, 플래시 셀의 데이터를 지우려면(즉, 플래시 셀의 상태를 '0' 상태에서 '1' 상태로 변경) 쓰기 동작과는 반대로 기판에 높은 전압($>20V$)을 인가하여 전하 트랩 층으로부터 전자를 제거함으로써 플래시 셀의 V_{th} 를 낮춘다. 플래시 셀에 저장된 데이터는 적절한 읽기 기준 전압 (V_{REF})을 WL 게이트에 인가함으로써 플래시 셀의 V_{th} 레벨을 감지하여 '0'과 '1'을 구분할 수 있다. 쓰기과 읽기 동작은 WL(또는 페이지) 단위로 수행되며 지우기 동작은 블록 단위로 수행된다.

NAND 플래시 메모리는 저장 용량을 증가시키기 위해서 2000년 중반 이후 하나의 플래시 셀에 여러 bit의 정보를 저장할 수 있는 멀티레벨 셀 기술을 적용하고 있다. 그림 3은 2^m 개의 V_{th} state를 사용하여 플래시 셀 내에 m 비트의 정보를 저장하는 m -bit 멀티레벨 NAND 플래시 메모리의 V_{th} 분포를 보여준다. 예를 들어 2-bit MLC NAND 플래시 메모리의 경우(즉, $m=2$ 인 경우) '00', '01', '10', '11'의 2 bit 정보를 서로 다른 V_{th} 분포를 갖는 E, P1, P2, 및 P3 4개의 state로 구분하여 저장한다. $m=3$ 인 경우는 TLC NAND 플래시 메모리이고 $m=4$ 인 경우가 QLC NAND 플래시 메모리이다.

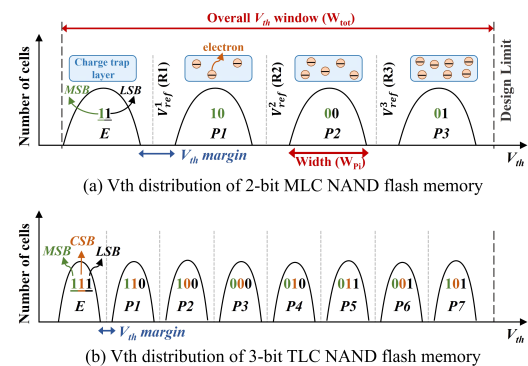


그림 3. m -bit 멀티레벨 NAND 플래시 메모리의 V_{th} 분포;
(a) 2-bit MLC NAND 플래시 메모리, (b) 3-bit TLC NAND 플래시 메모리
Fig. 3. V_{th} distributions of 2^m -state multi-level cell NAND flash memory; (a) 2-bit MLC NAND flash memory, (b) 3-bit TLC NAND flash memory.

2.2 NAND 플래시 메모리의 지우기 동작

Incremental step pulse erase (ISPE) 방식은 NAND 플래시 메모리의 지우기 동작을 수행하기 위한 현대

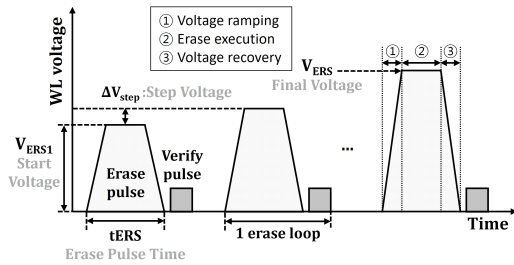


그림 4. Incremental Step Pulse Erase (ISPE) 기반 NAND 플래시 메모리의 지우기 동작

Fig. 4. Incremental Step Pulse Erase (ISPE) scheme in NAND flash memory erase operation.

SSD의 표준 방식이다¹⁶⁾. NAND 플래시 메모리의 수명과 저장된 데이터의 비트 에러를 효과적으로 관리하기 위해 NAND 플래시 칩은 단일의 매우 높은 전압 pulse(예를 들어, 20V)를 사용하여 지우기 동작을 수행하는 것이 아니라, 여러 개의 erase loop (일반적으로 5개 이하) 단위로 나누어 블록에 대한 지우기 동작을 반복적으로 수행한다. 개별 erase loop 수행 시 인가되는 erase 전압은 이전 erase loop보다 ΔV_{step} 만큼 높은 전압을 사용한다. 그림 4는 ISPE 기법이 동작하는 방식을 개략적으로 보여준다. 첫 번째 Erase 전압(V_{ERS1})을 인가하여 지우기 동작을 수행한 후 플래시 셀들의 상태를 검증하고(verify operation) 충분히 지워지지 않은 플래시 셀들에 대해서는 조금 더 높은 erase 전압($V_{ERS1} + \Delta V_{step}$)을 활용하여 지우기 동작을 반복적으로 수행한다. 이를 통해 플래시 셀이 과도하게 높은 erase 전압으로 인해 불필요하게 손상되는 것을 최소화하여 NAND 플래시 메모리 기반 저장장치의 수명을 보장한다¹⁷⁾. 하나의 erase pulse는 ① erase 전압이 원하는 값에 도달하는 voltage ramping 단계, ② 전압이 안정화 및 유지되는 erase execution 단계, ③ 다음 verify 동작을 위해 높은 erase 전압을 제거하는 voltage recovery 단계의 3단계로 구성된다. 따라서 하나의 erase loop는 (erase pulse + verify operation)으로 구성되며 지우기 동작의 응답시간은 블록의 모든 플래시 셀을 지우는 데 필요한 erase loop의 횟수에 비례하여 증가한다.

III. 기법 및 구현

3.1 기존 기법의 한계 및 핵심 아이디어

그림 5는 본 논문에서 제시하는 새로운 중단/재개 기법인 SSR의 핵심 아이디어를 기존에 제시되었던 기법들의 동작 원리와 비교하여 보여준다. 그림 5의 예시에서는 NAND 플래시 메모리의 블록 전체를 지우기

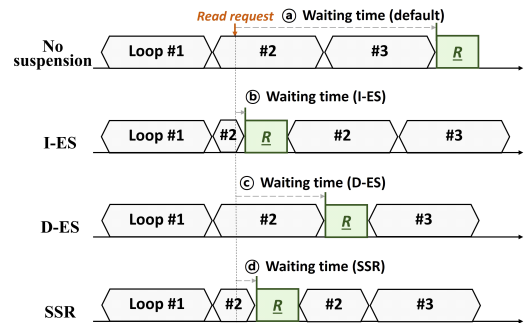


그림 5. NAND 플래시 메모리의 지우기 동작에 대한 중단/재개 기법

Fig. 5. Various suspend/resume techniques for NAND flash memory erase operation.

위해 3개의 erase loop가 수행되는 경우를 가정하였으며³⁾ 개별 기법에 대하여 읽기 요청의 예상 응답 지연 시간을 상대적인 비교를 통하여 보여준다.

중단/재개를 지원하지 않는 경우(Baseline) 지우기 동작과 호스트의 읽기 요청이 충돌하는 경우 읽기 요청은 지우기 동작이 완료될 때까지 긴 대기 시간을 경험한다 (그림 5의 ①). 이에 반해, I-ES 기법은 읽기 요청이 지우기 동작과 충돌 시 기존에 수행 중인 지우기 동작을 즉시 멈추고 호스트의 읽기 요청을 우선적으로 수행하며 이를 통해 100 μ s 수준의(기존의 지우기 동작을 중단하는 데 필요한 voltage recovery 시간) 최소한의 응답 시간 지연만을 겪는다 (그림 5의 ②).

하지만 서론에서 설명한 바와 같이 I-ES 기법은 높은 설계 비용 및 쓰기 동작의 기아 현상 등의 문제로 인해 상용 저장장치에 적용하기에는 한계가 있다. I-ES를 성공적으로 구현하기 위해서는 임의의 지점에서 erase pulse를 중단하고 다시 시작하는 기능이 필요하다. 따라서 임의의 길이의 erase pulse를 생성하고, 중단된 지우기 동작의 정확한 상태를 추적하고, 최종적으로 복구/재개하는 동작을 지원하기 위한 NAND 플래시 메모리 주변 회로(peripheral circuit)에 대한 추가적인 설계 비용이 발생한다. 특히 3D NAND 플래시 메모리의 도입으로 플래시 칩에서 NAND 플래시 셀이 차지하는 면적이 지속적으로 확장됨에 따라, NAND 플래시 메모리의 주변 회로를 추가하는 비용은 더욱더 커지고 있다^{18),19)}. 또한, 지우기 동작의 지연 및 이로 인한 쓰기 동작의 기아(starvation) 문제는 중단/재개 동작의 latency가 읽

3) 플래시 셀이 열화될수록 (즉, 플래시 셀이 쓰기/지우기 동작을 반복적으로 경험할수록) NAND 플래시 메모리를 지우는 데 필요한 erase loop 수는 증가한다 (지우기 동작의 응답시간 증가).

기 요청이 시스템에 인가되는 속도(incoming rate)보다 길 때 더욱 심각해진다. 예를 들어, PCIe Gen 3×4 NVMe 인터페이스의^[20] throughput이 3.2GB/s (즉, incoming read rate가 1.19us/4KB)이고 중단/재개 동작 수행을 위한 성능 overhead가 100us인 경우, 중단/재개 동작이 발생하는 동안 또 다른 읽기 요청이 도착할 수 있다. 이 경우, 지우기 동작은 즉시 다시 중단되므로 지우기 및 쓰기 동작은 심각한 기아 문제를 겪게 된다. NVMe 및 PCIe 사양에서는 쓰기 기아(starvation) 문제가 발생하면 호스트가 하드웨어와 소프트웨어를 모두 재설정해야 하며, 이는 시스템 성능에 악영향을 미치고 이러한 현상이 반복되면 결국 치명적인 시스템 장애로 이어질 수 있다. 마지막으로 가장 심각한 것은 제한 없이 인가되는 중단/재개 동작이 NAND 플래시 메모리의 신뢰성에 치명적인 문제를 일으킬 수 있다는 것이다. 중단 동작을 수행하기 위해서는 기판에 인가되었던 높은 지우기 동작 전압을 제거하는(그림 4의 voltage recovery) 구간이 추가로 필요하며 재개 동작 시에는 남은 지우기 동작을 완료하기 위하여 높은 지우기 동작 전압을 재인가하는(그림 4의 voltage ramping) 구간이 추가로 요구된다. 빈번한 중단/재개 기법이 인가되면 추가적인 voltage ramping/recovery 구간 동안에 인가되는 높은 동작 전압으로 인해 플래시 셀에 불필요한 손상을 유발하게 되며 결국 NAND 플래시 메모리 기반 저장장치의 수명은 급격하게 줄어들게 된다.

I-ES의 문제를 극복하기 위해서 D-ES는 지우기 동작과 호스트의 읽기 요청이 충돌하는 경우 현재 진행 중인 지우기 동작을 바로 중단하지 않고 해당 erase loop가 종료된 후 읽기 요청을 처리함으로써 지우기 동작 중에 인가되는 중단/재개 기법의 횟수를 제한할 수 있다. 예를 들어 지우기 동작을 완료하기 위해 3회의 erase loop가 필요한 경우 해당 지우기 동작 수행 중에 인가될 수 있는 중단/재개 기법의 횟수 또한 3회를 넘지 못하게 된다. 이를 통해 빈번한 중단/재개 기법으로 인한 NAND 플래시 메모리의 신뢰성 저하 문제를 방지하고 지우기 동작의 완료가 무한정 지연되는 문제도 해결할 수 있다. 하지만 읽기 요청에 대한 응답시간이 최대 5ms(1 erase loop 수행 시간, 그림 5의 ©)만큼 지연될 수 있어 꼬리 응답시간의 변동폭(QoS와 같은 저장장치의 성능지표) 개선에는 한계가 있다.

기존 기법의 한계를 극복하기 위해 본 논문에서는 3D NAND 플래시 메모리에 특화된 새로운 중단/재개 기법인 SSR (Smart Suspend Resume)을 제안한다. SSR은 1 erase loop 구간 안에서 신뢰성 열화 없이 수행할 수 있는 중단/재개 동작 횟수의 최대치를 기반으로

erase loop 구간에 N개의 safe point를 설정하고 호스트에서 읽기 요청이 인가되는 경우 지우기 동작을 바로 중단시키지 않고 해당 safe point까지 지우기 동작을 진행한 후 읽기 요청을 처리한다. SSR은 D-ES와 유사하게 중단/재개 동작의 발생 횟수를 제한함으로써 신뢰성 열화 문제를 효과적으로 회피함과 동시에 읽기 요청에 대한 응답시간 지연은 D-ES에 비해 평균 1/N배 개선시킨다(그림 5의 ④). 추가로 safe point는 off-line에서 측정된 디바이스 특성에 대한 모델링을 기반으로 결정되기 때문에(즉, 임의적으로 변하지 않음) 이를 구현하기 위한 추가적인 설계 비용도 최소화할 수 있다.

3.2 SSR 구현을 위한 디바이스 특성 모델링

SSR 구현의 핵심은 하나의 erase loop에 몇 개의 safe point를 할당할지를 결정하는 것이다. 만약 safe point의 개수가 N개라면 하나의 erase loop에 N번의 중단/재개 동작이 인가되더라도 NAND 플래시 메모리의 신뢰성은 열화되지 않는다는 것이 보장되어야 한다. 또한, safe point의 개수가 N으로 결정되면 읽기 요청의 꼬리 응답시간 지연의 최댓값은 (5ms/N)으로 제한된다. 따라서 safe point의 개수인 N을 결정하는 것이 저장장치의 성능과 신뢰성을 결정짓는 핵심 요소이다.

본 연구에서는 실제 최신 3D TLC NAND 플래시 칩 160개를 활용한 디바이스 특성 평가를 기반으로 SSR 구현의 핵심인 N값을 결정하였다. 평가 결과의 왜곡을 최소화하기 위해(평가 블록의 물리적인 위치에 따른 특성 차이를 반영하기 위해) 평가항목별로 서로 다른 물리적 위치에서 플래시 메모리 칩당 32개의 평가 블록을 선택하였으며 선택된 평가 블록의 모든 페이지를 테스트하였다. 그 결과 개별 평가항목별로 2,940,120 페이지(5,120 블록)를 평가하여 통계적으로 유의미한 결과를 얻을 수 있었다.

그림 6은 실제 NAND 플래시 칩을 탑재하여 평가하기 위해 자체적으로 제작한 평가 보드(NAND 컨트롤러 및 온도 컨트롤러 장착)를 보여준다. 이러한 평가 시스템을 사용하여 P/E 사이클(쓰기 동작과 지우기 동작의 반복) 및 데이터 보존 시간(retention time)과 같은 변수를 다양하게 변화시키면서 신뢰성 평가를 수행하였다. Retention을 포함한 신뢰성 평가는 블록에 특정 횟수의 P/E 사이클을 인가한 후 온도 30℃ 기준으로 0개월에서 12개월까지 방치하면서 1개월 단위로 비트 에러를 측정하면서 수행되었다. Retention 평가의 경우 평가 시간의 단축을 위하여 아레니우스 법칙을 활용한 가속 수명 평가를 수행하였다^[21]. 즉, Ea 1.1eV 기준으로 30℃ 1년의 retention 특성을 85℃ 13개월의 retention 특성을

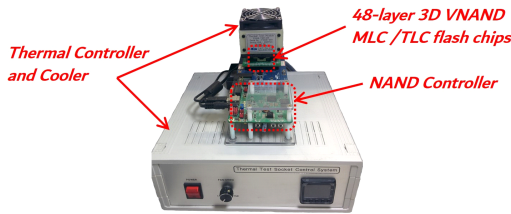


그림 6. 48단 3D NAND TLC 플래시 칩을 평가하기 위한 테스트 보드 구성

Fig. 6. Customized test board for evaluating real 3D TLC NAND flash memory chips.

통해 평가하였다. 모든 신뢰성 평가는 NAND 제조업체의 평가 표준인 JEDEC 기준에 따라 진행되었다^[22]. 또한, NAND 플래시 메모리의 비트 에러를 측정하기 위한 평가 패턴은 실제 상황을 모사하기 위하여 pseudo-random 패턴을 사용하였다^[23].

그림 7은 중단/재개 동작의 발생 횟수가 NAND 플래시 메모리의 신뢰성에 미치는 영향을 나타낸다. 정량적인 분석을 위해서 한 번의 지우기 동작 중 인가되는 중단/재개 동작의 발생 빈도를 증가시키면서(0회에서 100회까지) NAND 플래시 메모리의 비트 에러 발생률(Raw Bit Error Rate, RBER)의 변화를 확인하였다. 실험 결과의 정확성을 위해서 서로 다른 평가 블록을 선정하여 다양한 P/E 사이클을 인가한 상황에서 30°C 1년 retention 후의 RBER을 기준으로 평가를 진행하였다.

평가 결과를 통하여 세 가지의 핵심적인 관찰을 할 수 있었다. 첫 번째, 중단/재개 동작의 발생 빈도가 특정 횟수를 넘어서면 NAND 플래시 메모리의 RBER이 증가한다(즉, NAND 플래시 메모리 기반 저장장치의 수명이 감소). 두 번째, 중단/재개 동작의 발생 빈도가 높아질수록 NAND 플래시 메모리의 RBER은 급격히 증가한다(즉, 수명이 급격하게 감소). 마지막으로, NAND

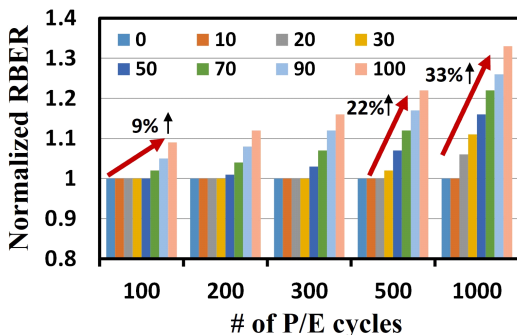


그림 7. 다양한 P/E 사이클 조건에서 중단/재개 동작의 횟수에 따른 비트 에러(RBER) 변화

Fig. 7. RBER increment with the number of suspend/resume operations under various P/E cycles.

플래시 메모리가 더 많은 P/E 사이클을 겪을수록(즉, 사용주기 후반부로 갈수록) 중단/재개 동작으로 인한 신뢰성 열화가 더 쉽게 발생한다.

3.3 SSR 기법 구현

중단/재개 기법과 신뢰성 간의 상관관계에 대한 NAND 플래시 메모리 특성 평가 결과를 바탕으로 SSR 구현을 위한 두 단계의 핵심 아이디어를 도출할 수 있다. 첫 번째, NAND 플래시 메모리의 모든 수명 주기에서 신뢰성의 열화를 방지하기 위해 지우기 동작을 구성하는 erase loop 1회에 대한 중단/재개 동작을 10회로 제한한다. 다시 말해서 erase loop 1회에 대해서 10개의 safe point를 설정할 수 있다. 허용 가능한 safe point의 개수는 최악의 동작 조건(예를 들어 수명이 거의 다한 낸드 플래시 메모리가 1년 동안 데이터를 저장한 후 복구하는 상황)을 기준으로 중단/재개 기법으로 인한 추가 RBER의 발생을 억제할 수 있는 수준으로 결정된다. 상기 접근 방법은 NAND 플래시 메모리의 모든 수명 주기에 대하여 safe point의 개수를 정적으로 할당함으로써 낮은 설계 비용으로 SSR을 구현할 수 있지만, 저장장치의 꼬리 응답시간은 최대 5ms/10(safe point의 개수), 500 μ s에 이르게 된다. 본 논문에서는 해당 기법을 static SSR, SSR(s), 라고 명명한다. 그림 8의 상단 그림에서 확인할 수 있듯이 SSR(s)는 모든 수명 주기에 대해서(즉, P/E cycle에 무관하게) erase loop 1회에 대한 중단/재개 동작을 허용하는 safe point 횟수가 10회로 고정되어 있으며 이로 인한 응답시간 지연도 최대 500 μ s로 고정된다.

두 번째 단계로 SSR 기법의 꼬리 응답시간 개선 효과를 극대화하기 위하여 디바이스 평가 결과를 활용한 추가 최적화 방법을 제안한다. 그림 7에서 보듯이 신뢰성 열화 없이 허용 가능한 중단/재개 동작의 최대 횟수

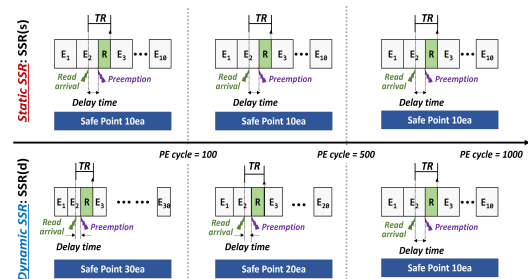


그림 8. 다양한 P/E 사이클 수명 조건에 따른 safe point 할당 방식: (a) static SSR, (b) dynamic SSR

Fig. 8. Allocation the number of safe points in the erase operation with varying P/E cycles; (a) static SSR, (b) dynamic SSR.

는 P/E 사이클 횟수에 따라 다르게 결정될 수 있다. 예를 들어 낸드 플래시 메모리의 열화 정도가 낮을 경우 (즉, P/E cycle을 경험한 횟수가 낮은 경우)에는 더 많은 safe point를 설정하더라도 RBER 증가가 억제되지만 낸드 플래시 메모리의 열화가 진행되면서 수명 주기 후반으로 갈수록(즉, P/E cycle을 경험한 횟수가 많은 경우로 end-of-life, EOL 상황) RBER 증가를 억제하기 위해서는 더 적은 수의 safe point를 할당해야 한다. 따라서 NAND 플래시 메모리의 열화 상태에 따라(얼마나 많은 양의 P/E 사이클을 겪었는지에 따라) 지우기 동작 중에 중단/재개 작업을 허용하는 safe point의 개수를 동적으로 할당하면 저장장치의 수명 주기 대부분에 대하여 SSR(s)에 비해 더 많은 수의 safe point를 설정할 수 있다. 본 논문에서는 해당 기법을 dynamic SSR, SSR(d), 라고 명명한다. 그림 8의 상단 그림에서 확인할 수 있듯이 SSR(d)는 낸드 플래시 메모리가 경험한 P/E cycle 횟수에 따라 safe point 개수를 동적으로 변경하며 할당한다. 예를 들어 P/E cycle이 100회로 열화 정도가 낮은 경우에는 30개의 safe point를 할당하지만, 열화 정도가 높은 P/E cycle 1,000회의 경우에는 safe point를 10개만 인가하여 신뢰성 열화를 방지한다. 이를 통해 대부분의 수명 주기에 대하여 SSR(s)에서 허용된 10개의 safe point보다 더 많은 safe point를 할당할 수 있다. 읽기 동작의 꼬리 응답시간은 safe point의 개수에 반비례하여 개선되므로 SSR(d)는 저장장치의 수명 열화 없이 저장장치의 꼬리 응답시간 개선을 극대화할 수 있다.

IV. 평가 결과

4.1 평가 방법

본 논문에서는 저장장치 연구에서 가장 널리 사용되는 시뮬레이터인 MQSim을 활용하여 다양한 조건에서 성능 평가를 진행하였다. MQSim을 확장하여 다양한 중단/재개 기법(Baseline, I-ES, D-ES, 그리고 SSR)을 지원하는 저장장치를 구현하여 SSR의 효과를 정량적으로 비교 검증하였다.

개별 기법에 대하여 Flexible I/O Tester (FIO) 와 TPC-C 벤치마크를 활용하여 읽기와 쓰기 응답 시간에 대한 시스템 성능 평가를 진행하였다. FIO^[26]는 다양한 읽기/쓰기 비율을 갖는 4KB 랜덤 액세스 워크로드를 생성하며 SSD의 latency 성능 평가에 널리 활용된다. TPC-C^[27]는 온라인 트랜잭션 처리 시스템의 성능 평가에 널리 사용되는 벤치마크이다. 또한, 우리는 모든 평가를 시작하기 전에 over-provisioning (OP) 영역을 포

표 1. SSD를 모델링하는 데 사용되는 매개 변수
Table 1. Parameters used to model our SSD.

NVMe SSD Device	
NAND Configurations	4 channels, 4 chips/channel, 1 die/chip
FTL Schemes	Page Mapping, Preemptible GC ^[24]
Over-provisioning	7% ^[25]
NAND Flash Memory	
128GB die capacity, 2 planes per die, 683 blocks per plane, 768 pages per block, 16KB page	
NAND Latency	
Read: 40us, Program: 660us, Block Erase: 5ms per step (5 steps), Erase Suspension Penalty: 100us	

함한 SSD의 모든 저장 공간에 대하여 쓰기 동작을 수행하여 동일한 사전 상태를 설정하였다^[28]. 이러한 사전 조건은 최악의 조건에서 발생할 수 있는 읽기/쓰기 동작의 latency를 평가하는 데 효과적이다. 표 1은 SSD를 모델링 하기 위해 사용된 NAND 플래시 메모리 관련 파라미터들을 보여준다.

4.2 평가 결과

그림 9는 대기열 깊이(queue depth) 16에서 4KB 랜덤 읽기(70%)와 쓰기(30%)의 혼합을 생성하는 micro-benchmark를 사용하여 다양한 중단/재개 동작의 정책을 비교 평가한 결과이다. 그림 9(a)에서 볼 수 있듯이 Baseline에서는 지우기 동작이 발생하면 이후의 읽기 요청들은 지우기 동작이 완료될 때까지 지연된다. 이에 반해, I-ES는 읽기 요청이 인가되는 즉시 지우기 동작을 중단하고 읽기 요청을 처리하기 때문에 응답시간만을 고려한다면 100 μ s 수준의 응답시간 지연으로 가장 유리하다. 지우기 동작을 중단하는 횟수를 제한하는 D-ES, SSR(s), 및 SSR(d)의 경우 제한 횟수를 더 많이 허용할수록 낮은 읽기 응답시간 지연을 보이는 것을 확인할 수 있다.

그림 9(b)는 중단/재개 동작의 정책에 따른 쓰기 동작의 꼬리 응답시간을 비교해서 보여준다. I-ES가 읽기 응답시간 측면에서 가장 유리하지만 지우기 동작 또는 쓰기 동작이 긴 시간 동안 지연될 수 있기 때문에 쓰기 꼬리 응답시간 측면에서는 최악의 결과를 보인다. 이에 반해, D-ES, SSR(s), 및 SSR(d)에서는 읽기 작업이 요청되더라도 일정 시간 동안 지우기 동작을 수행하기 때문에 지우기/쓰기 동작의 기아 문제를 효과적으로 해결할 수 있으며 이에 따라 쓰기 꼬리 응답시간 측면에서 장점을 갖는다. 그리고 읽기 요청 이후에 중단되지 않고

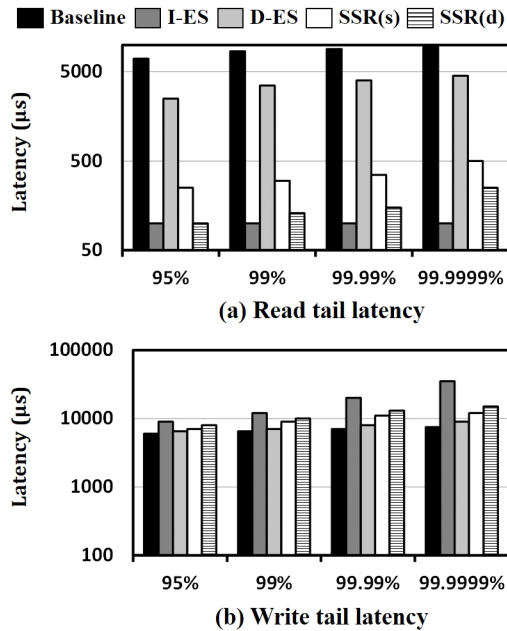


그림 9. 4KB 랜덤 읽기(70%) 및 쓰기(30%) 워크로드에 대한 (a) 읽기 꼬리 응답시간, (b) 쓰기 꼬리 응답시간
Fig. 9. Tail latency on 4KB random reads (70%) and writes (30%) workload: (a) Read tail latency and (b) Write tail latency.

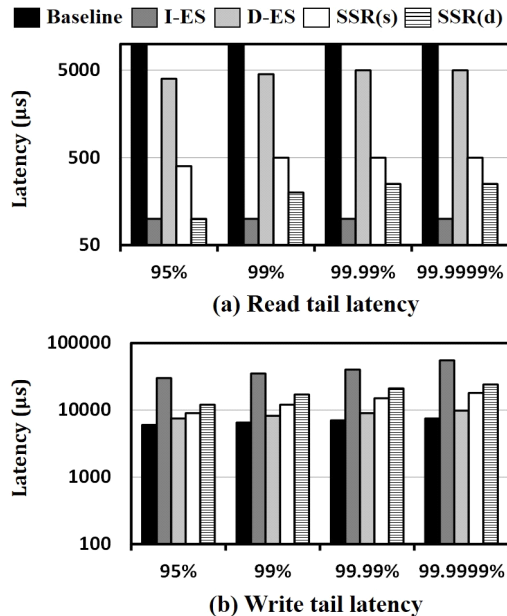


그림 10. TPC-C 워크로드에 대한 (a) 읽기 꼬리 응답 시간, (b) 쓰기 꼬리 응답 시간
Fig. 10. Tail latency on TPC-C workload: (a) Read tail latency and (b) Write tail latency.

수행되는 지우기 동작의 허용 시간의 크기가 클수록 (I-ES < SSR(s) < SSR(d) < D-ES) 쓰기 동작의 꼬리 응답시간이 개선되는 것을 확인할 수 있다.

그림 10은 TPC-C 워크로드에 대한 읽기/쓰기 동작의 꼬리 응답시간 평가 결과이다. TPC-C의 평가 결과는 FIO를 활용한 4KB 랜덤 패턴에 대한 평가와 비교 시 그 경향은 유사하지만 많은 읽기 동작이 지우기 동작 뒤에 대기함에 따라 읽기 요청에 지우기 동작의 긴 응답시간이 노출될 확률이 높기 때문에 상대적으로 긴 꼬리 응답시간의 지연을 보여준다.

결과적으로 읽기 응답시간과 쓰기 응답시간은 서로 trade-off의 보상관계를 갖는 것을 알 수 있다. 따라서 본 논문의 기법은 SSR(s)와 SSR(d)를 워크로드의 특성(예를 들어, 읽기 성능이 중요한 응용인지 쓰기 성능이 중요한 응용인지) 따라 선택적으로 적용함으로써 읽기 응답시간과 쓰기 응답시간 모두를 균형적으로 최적화할 수 있는 해결책을 제시한다.

V. Related Work

Garbage Collection Optimization. 꼬리 응답시간에 대한 GC의 영향을 완화하기 위한 몇 가지 선행 연구가 있다^[6,7,29,30]. 이러한 최적화는 읽기 응답시간에 대한 GC의 영향을 효과적으로 감소시킬 수 있다. 하지만, 최적화된 GC에서 더 중요해지는 읽기 동작의 꼬리 응답시간에 대한 긴 지우기 동작의 영향은 다루지 않는다. 따라서 이러한 선행 기술들은 우리가 제시한 기법과 상호 보완적으로 적용할 수 있다.

I/O Scheduling Optimization. 꼬리 응답시간을 최적화하는 또 다른 접근 방법으로 읽기/쓰기 요청을 스케줄링하는 기법들이 연구되었다^[24,32-34]. 이러한 제안들은 이용 가능한 다수의 디바이스가 존재할 때 효과적이다. 반면에 본 논문의 기법은 개별 NAND 플래시 메모리 칩 하나에 대한 꼬리 응답시간의 개선에 중점을 두고 있다. 따라서 기존 최적화 기법들은 본 논문이 제시하는 기법과 충돌 없이 함께 사용될 수 있다.

VI. 결 론

본 논문에서는 지우기 동작이 읽기 동작의 꼬리 응답시간 지연에 미치는 영향을 제한하기 위해 최신 플래시 기반 저장장치에서 일반적으로 사용되는 ISPE 메커니즘 기반의 새로운 중단/재개 기법 SSR을 제안한다. SSR은 최신 3D TLC NAND 플래시 칩 160개를 활용한 평가를 통해 구축한 디바이스 특성 모델링을 바탕으로

로 단일 erase loop 동안에 신뢰성 열화 없이 허용 가능한 중단/재개 기법의 횟수를 확인하고 이를 기반으로 중단/재개 기법을 지원할 수 있는 제한된 safe point를 설정한다. Safe point에서만 중단/재개 동작을 허용함으로써 플래시 인터페이스에 대한 약간의 확장만으로 신뢰성 열화와 지우기/쓰기 동작의 기아 문제를 효과적으로 회피할 수 있다. 추가로 NAND 플래시 메모리의 상태에 따라 safe point 횟수를 동적으로 할당하여 저장장치 시스템의 읽기 동작의 꼬리 응답시간 지연을 최소화할 수 있다. MQSim을 활용한 시뮬레이터 기반 평가를 통해 SSR은 기존 기법 대비 99.99%, 99.9999% 꼬리 응답시간을 최대 10배 개선시킬 수 있음을 확인하였다.

본 연구의 기법은 NAND 플래시 메모리 기반 저장장치의 잠재력을 최대한 활용하여 응답속도에 민감한 다양한 차세대 응용들을 실행하는 데 필수적인 솔루션이 될 수 있다.

References

- [1] M. Hao, G. Soundararajan, D. Kenchammana, A. Chien, and S. Gunawi, "The tail at store: A revelation from millions of hours of disk and SSD deployments," in *Proc. USENIX Conf. FAST*, pp. 263-275, Santa Clara, CA, USA, Feb. 2016.
- [2] W. Cheong, C. Yoon, S. Woo, K. Han, D. Kim, C. Lee, Y. Choi, S. Kim, D. Kang, G. Yu, J. Kim, J. Park, K. Song, K. Park, S. Cho, H. Oh, D. Lee, J. Choi, and J. Jeong, "A flash memory controller for 15 μ s ultra-low-latency SSD using high-speed 3D NAND flash with 3 μ s read time," in *Proc. IEEE ISSCC*, San Francisco, CA, USA, 2018.
- [3] K. Alizadeh, I. Mirzadeh, D. Belenko, K. Khatamifard, M. Cho, C. Del Mundo, M. Rastegari, and M. Farajtabar, "LLM in a flash: Efficient large language model inference with limited memory," *arXiv preprint arXiv:2312.11514*, 2013.
- [4] M. Chun, J. Lee, M. Kim, J. Park, and J. Kim, "RiF: Improving read performance of modern SSDs using an on-die early-retry engine," in *Proc. IEEE Int. Symp. HPCA*, pp. 643-656, Edinburgh, UK, 2024.
- [5] Samsung, *Quality of Service (QoS) Quality of Service to offer predictable response times (2014)*, 01, 03, 2025, from https://download.semiconductor.samsung.com/resources/handling-guide/Samsung_SSD_845DC_06_Quality_of_Service_QoS.pdf.
- [6] W. Kang, D. Shin, and S. Yoo, "Reinforcement learning-assisted garbage collection to mitigate long-tail latency in SSD," *ACM TECS*, vol. 16, no. 5s, pp. 1-20, 2017.
(<https://doi.org/10.1145/3126537>)
- [7] S. Yan, H. Li, M. Hao, M. Tong, S. Sundararaman, A. Chien, and H. Gunawi, "Tiny-tail flash: Near-perfect elimination of garbage collection tail latencies in NAND SSDs," *ACM TOS*, vol. 13, no. 3, pp. 1-26, 2017.
(<https://doi.org/10.1145/3121133>)
- [8] Z. Du, Z. Dong, K. You, X. Jia, Y. Tian, Y. Wang, Z. Yang, X. Fu, F. Liu, Q. Wang, L. Jin, and Z. Huo, "A novel program suspend scheme for improving the reliability of 3D NAND flash memory," *IEEE J. EDS*, vol. 10, pp. 98-103, 2022.
(<https://doi.org/10.1109/JEDS.2022.3140949>)
- [9] C. Zambelli, L. Zuolo, A. Aldarese, S. Scommegna, R. Micheloni, and P. Olivo, "Assessing the role of program suspend operation in 3D NAND flash based solid state drives," *Electr.*, vol. 10, no. 12, pp. 1-18, 2021.
(<https://doi.org/10.3390/electronics10121394>)
- [10] G. Wu, P. Huang, and X. He, "Reducing SSD access latency via NAND flash program and erase suspension," *JS4*, vol. 60, no. 4, pp. 345-356, 2014.
(<https://doi.org/10.1016/j.sysarc.2013.12.002>)
- [11] S. Kim, J. Bae, H. Jang, W. Jin, J. Gong, S. Lee, T. Ham, and J. Lee, "Practical erase suspension for modern low-latency SSDs," in *Proc. USENIX ATC*, pp. 813-820, Renton, WA, USA, 2019.
- [12] R. Micheloni, L. Crippa, and A. Marelli, *Inside NAND Flash Memories*, Springer, 2010.
(<https://doi.org/10.1007/978-90-481-9431-5>)

- [13] D. Hong, M. Kim, G. Cho, D. Lee, and J. Kim, "GuardedErase: Extending SSD lifetimes by protecting weak wordlines," in *Proc. USENIX Conf. FAST*, pp. 133-146, Santa Clara, CA, USA, Feb. 2022.
- [14] A. Tavakkol, G. Juan, M. Sadrosadati, S. Ghose, and O. Mutlu, "MQSim: A framework for enabling realistic studies of modern multi-wueue SSD devices," in *Proc. USENIX Conf. FAST*, pp. 49-66, Santa Clara, CA, USA, Feb. 2018.
- [15] J. Maserjian and N. Zamani, "Behavior of the Si/SiO₂ interface observed by fowler-nordheim tunneling," *JAP*, vol. 53, no. 1, pp. 559-567, 1982.
(<https://doi.org/10.1063/1.329919>)
- [16] D. Lee, S. Cho, B. Kang, S. Park, B. Park, M. Cho, K. Ahn, Y. Yang, and S. Park, "The operation algorithm for improving the reliability of TLC (triple level cell) NAND flash characteristics," in *Proc. IEEE IMW*, pp. 1-2, Monterey, CA, USA, May 2011.
- [17] J. Jeong, S. Hahn, S. Lee, and J. Kim, "Lifetime improvement of NAND flash based storage systems using dynamic program and erase scaling," in *Proc. USENIX Conf. FAST*, pp. 61-74, Santa Clara, CA, USA, Feb. 2014.
- [18] J. Lee, J. Jang, J. Lim, Y. Shin, K. Lee, and E. Jung, "A new ruler on the storage market: 3D-NAND flash for high-density memory and its technology evolutions and challenges on the future," in *Proc. IEEE IEDM*, pp. 11.2.1-11.2.4, San Francisco, California, USA, Dec. 2016.
- [19] B. Park, S. Cho, M. Park, S. Park, Y. Lee, M. Cho, K. Ahn, G. Bae, and S. Park, "Challenges and limitations of NAND flash memory devices based on floating gates," in *Proc. IEEE ISCAS*, pp. 420-423, Seoul, Korea, May 2012.
- [20] *NVM express[®] Base Specification (2024)*, 01, 03, 2025, from <https://nvmexpress.org/specifications/>
- [21] S. Arrhenius, "Über die dissociationswärme und den einfluss der temperatur auf den dissociationsgrad der elektrolyte," *Zeitschrift für Physikalische Chemie*, pp. 96-116, 1889.
(<https://doi.org/10.1515/zpch-1889-0408>)
- [22] JEDEC, *Stress-test-driven qualification of integrated circuits (JEDEC jesd47) (2010)*, 01, 03, 2025, from <https://www.jedec.org>
- [23] M. Favalli, C. Zambelli, A. Marelli, R. Micheloni, and P. Olivo, "A scalable bidimensional randomization scheme for TLC 3D NAND flash memories," *Micromachines*, vol. 12, no. 7, pp. 759-782, 2021.
(<https://doi.org/10.3390/mi12070759>)
- [24] J. Lee, Y. Kim, G. Shipman, S. Oral, and J. Kim, "Preemptible I/O scheduling of garbage collection for solid state drives," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Syst. (TCAD)*, vol. 32, no. 2, pp. 247-260, 2013.
(<https://doi.org/10.1109/TCAD.2012.2227479>)
- [25] S. Kent, "Understanding SSD overprovisioning," in *Proc. Flash Memory Summit (FMS)*, Santa Clara Francisco, CA, USA, Aug. 2013.
- [26] *Flexible IO Tester*, 01, 03, 2025, from <https://github.com/axboe/fio>
- [27] *TPC-C Benchmark*, 01, 03, 2025, from <http://tpcc.org/tpcc/>
- [28] *SNIA Solid State Storage Performance Test Specification*, 01, 03, 2025 from https://www.snia.org/sites/default/files/HoEasen_SNIA_Solid_State_Storage_Per_Test_1_0.pdf
- [29] W. Choi, M. Jung, M. Kandemir, and C. Das, "Parallelizing garbage collection with I/O to improve flash resource utilization," in *Proc. Int. Symp. HPDC*, pp. 243-254, Tempe, AZ, USA, Jun. 2018.
- [30] J. Kim, D. Lee, and S. Noh, "Towards SLO complying SSDs through OPS isolation," in *Proc. USENIX Conf. FAST*, pp. 183-189, Santa Clara Francisco, CA, USA, Feb. 2015.
- [31] G. Zhu, J. Han, S. Lee, E. Lee, and Y. Son, "A parallel garbage collection scheme for SSDs," in *Proc. KCC*, pp. 1179-1181, Oline, Jul. 2020.
- [32] S. Wu, W. Zhu, G. Liu, H. Jiang, and B.

Mao, "GC-aware request steering with improved performance and reliability for SSD-based RAID," in *Proc. IEEE IPDPS*, pp. 296-305, Vancouver, BC, Canada, May 2018.

- [33] J. Zhang, A. Sivasubramaniam, Q. Wang, A. Riska, and E. Riedel, "Storage performance virtualization via throughput and latency control," in *Proc. IEEE Int. Symp. MASCOTS*, pp. 135-142, Atlanta, GA, USA, Sep. 2005.
- [34] Q. Zhang, D. Feng, F. Wang, and Y. Xie, "An efficient, QoS-aware I/O scheduler for solid state drive," in *Proc. IEEE Int. Conf. HPCC*, pp. 1408-1415, Zhangjiajie, China, Nov. 2013.

김 범 준 (Beomjun Kim)



2021년 2월 : 경북대학교 컴퓨터 학부 졸업
 2022년 8월 : 경북대학교 컴퓨터 학부 석사
 2022년 9월~현재 : 경북대학교 컴퓨터학부 박사과정
 <관심분야> SSD, 낸드 플래시 메모리, 비휘발성 메모리

[ORCID:0000-0003-1211-5641]

서 경 섭 (Gyeongseob Seo)



2023년 2월 : 경북대학교 컴퓨터 학부 졸업
 2023년 3월~현재 : 경북대학교 컴퓨터학부 석사과정
 <관심분야> 낸드 플래시 메모리, 임베디드 시스템 및 소프트웨어

[ORCID:0009-0008-9901-9874]

김 명 석 (Myungsuk Kim)



2003년 : 연세대학교 전기전자공학과 학사
 2005년 : 연세대학교 전기전자공학과 석사
 2020년 : 서울대학교 컴퓨터공학과 박사
 2005년~2020년 : 삼성전자 메모리사업부 Principal Engineer

2021년~2024년 : 경북대학교 컴퓨터학부 조교수
 2025년~현재 : 경북대학교 컴퓨터학부 부교수
 <관심분야> 메모리 및 저장장치 시스템, 임베디드 시스템 및 소프트웨어 등

[ORCID:0000-0002-8667-3198]