

# Guarded Erase: Extending SSD Lifetimes by Protecting Weak Wordlines

## Abstract

- 3D NAND의 손상을 줄이기 위해 수직으로 WL을 쌓는 기법을 도입하였으며 이는 WL에 각각 다른 주명을 유발한다.
- 이를 해결하기 위해 우리는 system-level block scheme인 Guarded Erase를 제안한다.
- 이는 불러울 삽입할 때 low-stress erase mode와 normal erase mode 두 가지를 선택할 수 있으며 이는 주명 확장에 큰 영향을 미친다.
- 우리는 Intel SSD에 이를 구현하였으며 SSD 수명은 21% 향상되었고 성능 차이는 미미하게 발생하였다.

## Introduction

- WL을 수직으로 쌓았기 때문에 기존 2D-NAND Flash Chip보다 3D 구조에서 용량이 크게 증가하는 경향이 있다.

표 1을 보거나 강의 block 크기는 크게 증가하였으며 페이지 크기는 이와 다르게 크게 증가하지 않았기 때문에 3D 구조에는 물론은 WL을 포함할 수 밖에 없다.



- 3D NAND는 etching 공정을 사용하기 때문에 WL의 top에서 bottom 까지 큰 변동성을 가진다.

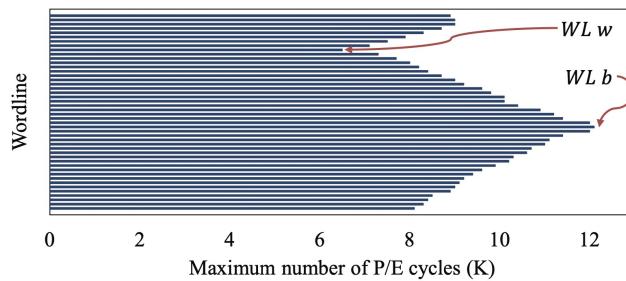


Figure 1: Per-WL variations on the maximum number of P/E cycles.

- 공정의 변동성은 WL마다 다른 수명을 가져온다. 그림 1을 보면 이에 대한 명확성이 잘 나타난다.
- 결국 block의 수명은 제일 약한 WL로 결정된다. 이러한 block의 관리는 BBM (Block Block Management)에 의해 관리된다.

- Coarse-grained BBM의 단점을 해결하기 위해 BDM (Bind page Management)을 제안한다.
- BDM은 SSD 용량을 줄이며 이는 I/O 성능을 크게 감소시킨다.
- 해결책 중 하나는 가장 약한 WL을 더 사용하는 방법이며 이는 2D 메모리 기술에서 많이 사용되었던 기술이다.
- NAND Flash의 모든 program 블록은 Erase 단위에서 주된 원인이 되어 발생한다.
- Guard Erase는 효과적인 지우기 방법으로 불러울 수명을 효과적으로 연장 시킬 수 있다.

$$TBWS = \frac{C_S \times \text{MAX}_{P/E}}{WAF},$$

↑ total write amplification factor  
of an FTL including GC and Wear leveling

the capacity of SSD

Maximum number of P/E cycle

the total amount of written data in term bytes

- Guard Erase 2] 저스트레이스 모드는 지우기 전파를 재조정하는 것을 목표로 한다.
- 저스트레이스 모드에서 마모도가 33% 정도이며 이를 위해 long-FTL을 제안한다.
- MQ-SIM에서 이를 실현하였으나 21%, SSD 수명이 증가하였으며 오직 3% 성능 저하만 나타나았다.

## Background .

### 2.1 Overview of 3D NAND Flash Memory

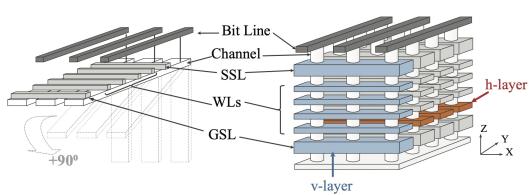


Figure 2: Differences between 2D NAND flash and 3D NAND flash [9].

- 3D NAND flash memory로 발전하면서 SSD의 용량은 크게 증가하였지만 신뢰성 문제에서는 큰 차이를 보여주고 있다.
- SSD에 의해 목표로 하는 WL이 선택되어 그쪽에 따라 수평적으로 쌓여 있다.

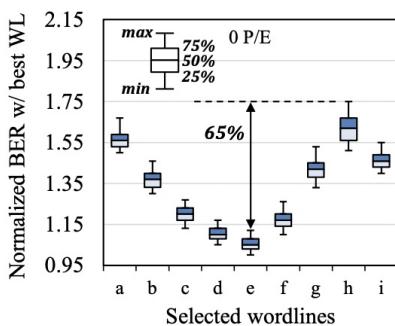
### 2.2 Reliability Management in NAND Flash Memory

- 많은 erase 작업은 tunnel oxide 양극에 더 많은 전자를 trapping 시기며 이를 의도된 전극 양극에 도달하는 것을 어렵게 만들고 BER을 악화 시킨다.
- 이를 방지하기 위해 회자기의 1/4 cycle를 제한하여 FTL의 능력을 초과하게 되면 데이터 손실을 일으키게 된다.
- 마모 이외에도 제조 프로세스와 같은 원인으로 다양한 이유가 발생할 수 있다.
- Bad block은 FTL 내부에 바로 관리된다.

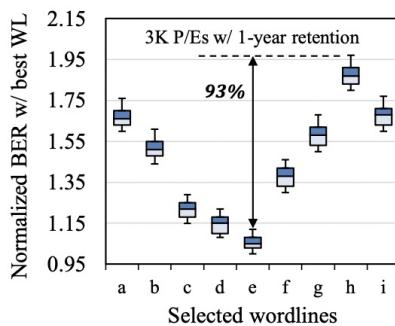
## 3 Reliability Characterization of 3D blocks

### 3.1 Large Reliability Variations Among WLs

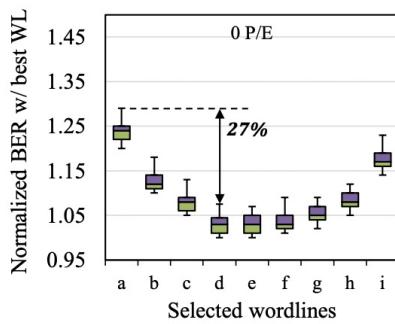
- 2D NAND에서 WL의 물리적 위치에 따라 신뢰성이 다르다는 것이 알려져있고 이는 3D 메모리에서 더욱 심각이다.



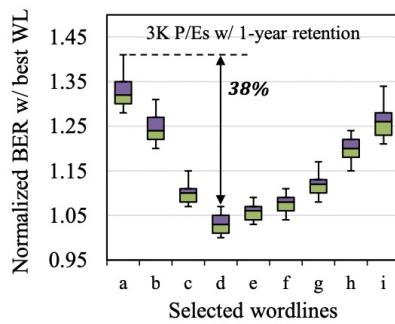
(a) Per-WL BER of 3D flash.



(b) Per-WL BER of 3D flash.



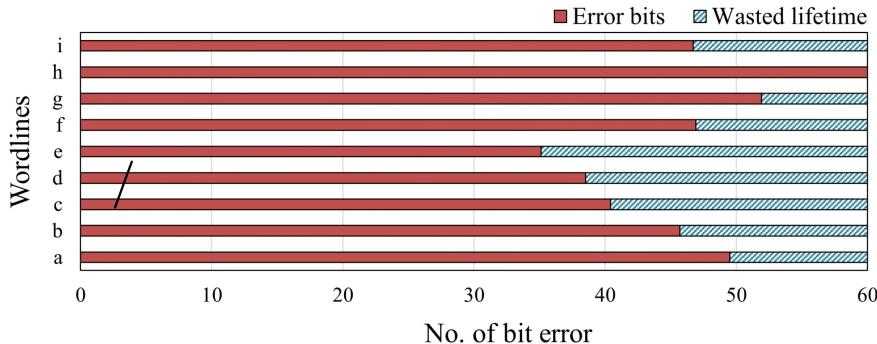
(c) Per-WL BER of 2D flash.



(d) Per-WL BER of 2D flash.

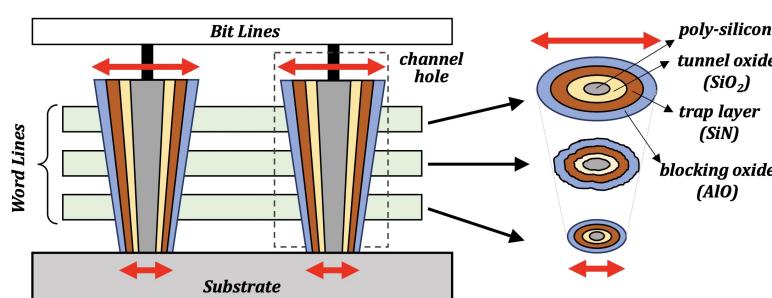
Figure 3: Per-WL BER variations in 3D NAND flash and 2D NAND flash.

- 0 cycle에 | Best WL과 Worst WL의 BER 차이는 크게 65%를 보인다.
- Block이 age 되었을 때 이 차이는 크게 93% 증가하게 된다.
- 2D NAND Flash는 3D NAND보다 이 차이가 작지만 유사한 경향을 보여주고 있다.



- 최악의 WL이 60c 능력을 초과했더라도 나머지 WL은 상당히 신뢰할 수가 있다.
- 이는 상당한 Block 수행을 감소시키는 원인으로 작용한다.

Figure 4: Inter-WL BER variations when the worst WL becomes bad.



(a) A cross-sectional view.      (b) A top-down view.

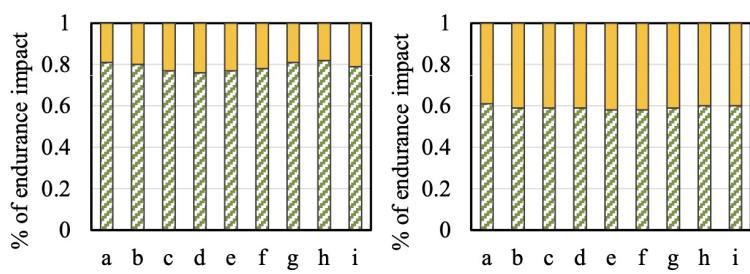
Figure 5: Inter-WL variations in 3D NAND flash memory.

### 3.2 Erase Stress on Flash Reliability

Table 2: A summary of three operation sequences.

Sequence type	Operations order
Modified Base Sequence	P → dummy P → E → dummy E
Erase-only Sequence	dummy P → E → dummy E
Program-only Sequence	dummy E → P → dummy P

P : program, E : erase



(a) 3D NAND flash.      (b) 2D NAND flash.

Figure 6: Endurance impact of erase and program operations.

- 3D NAND는 이상적으로 일정한 크기의 월등한 구조를 가지고 있지만 하지만 결정 특성상 불균형을 피할 수가 없으며 이는 신뢰성 향상에서 높이에 따른 다른 결과를 보여줄 수 밖에 없다.

- Endurance 흐름이 3D NAND에서 더욱 크게 나타나는데 이는 총 스트리스의 80%를 차지하는 결과를 보여준다.
- 이 영향은 3D NAND가 2D보다 33% 더 높은 비율을 보여준다.
- 전자적으로 erase가 program보다 더 많은 영향을 끼치는 것을 보여준다.

### 3.3 Erase Stress Reduction

- erase stress를 줄이는 것은 tunnel oxide의 대비치를 줄일 수 있어 필수적이다.
- 제거 전압이 증가하면 tunnel oxide의 속성이 금속하게 증가한다.

### 4 Guarded Erase : Mechanism and Modes

#### 4.1 Basic Idea

- 각 P/E cycle은 각각 스트레스를 일으키기 때문에 봇 쓰기에 따른 low-stress mode와 normal erase mode 두 가지를 제작한다.
- lower erase voltage mode로 데이터를 사용 경우 WL의 주입이 증가한다.

WL 9	1   2   3   4   5   6   7   8   9   10
WL 8	1   2   3   4   5   6   7   8   9   10
WL 7	1   2   3   4   5   6   7   8   9   10
WL 6	1   2   3   4   5   6   7   8   9   10
WL 5	1   2   3   4   5   6   7   8   9   10
WL 4	1   2   3   4   5   6   7   8   9   10
WL 3	1   2   3   4   5   6   7   8   9   10
WL 2	1   2   3   4   5   6   7   8   9   10
WL 1	1   2   3   4   5   6   7   8   9   10
WL 0	1   2   3   4   5   6   7   8   9   10

Increasing BER values → BER threshold

1   2   3   4   5   7   9   10   11   13   14
1   2   3   4   5   6   7   8   9   10   11   13   14
1   2   3   4   5   6   7   8   9   10   11   12   13   14
1   2   3   4   5   6   7   8   9   10   11   12   13   14
1   2   3   4   5   6   7   8   9   10   11   12   13   14
1   2   3   4   5   6   7   8   9   10   11   12   13   14
1   2   3   4   5   6   7   8   9   10   11   12   13   14
1   2   3   4   5   6   7   8   9   10   11   12   13   14
1   2   3   4   5   6   7   8   9   10   11   12   13   14
1   2   3   4   5   6   7   8   9   10   11   12   13   14

Increasing BER values → BER threshold

(a) Using the normal erase mode. (b) Using the gErase mode.

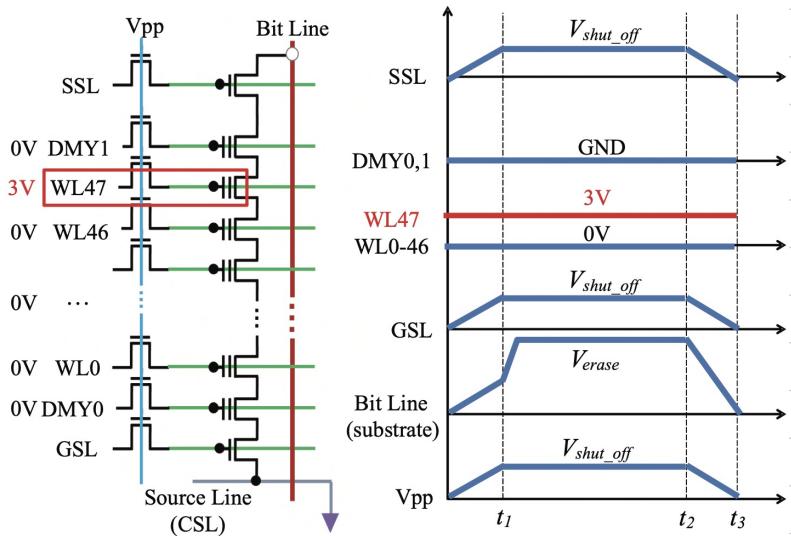
- 그림 7(a)를 보면 정상적인 WL이 BER을 초과하는 경우에 BBM에 의해 해방 block은 폐기 처리되게 되고 있다. 반면 세밀하게 block을 관리하게 되면서 낭비되는 용량은 줄일 수가 있다.
- 그림 7(b)는 약한 WL을 white box mode로 처리함으로써 low stress가 발생하는데, 이는 강한 WL의 14번의 cycle을 가질 수 있게 해준다.

Figure 7: Per-WL BER changes under the normal block erase and the gErase block erase.

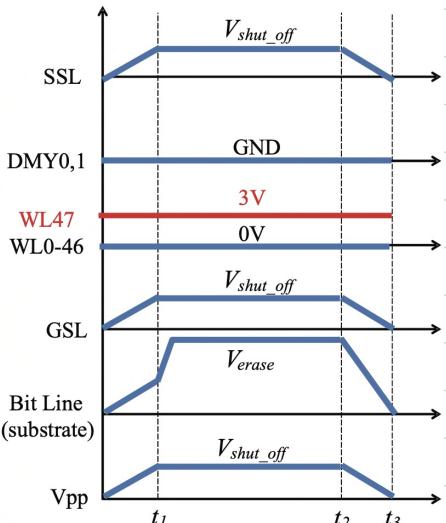
### 4.2 Per-WL Low-Stress Erase Mode

#### 4.2.1 Implementation

- two-options ① reduce the erase time → It's very hard.  
 ② reduce the erase voltage → our choice.  
 (test mode command을 적용하였다)



(a) Differential voltage driving to control gates.



- low-stress mode는 약한 WL을 약한 전압으로 인가하여 기판과의 접촉차를 크게 만든다.

Figure 8: An implementation of the low-stress erase mode.

## 4.2 Stress Mitigation Effect

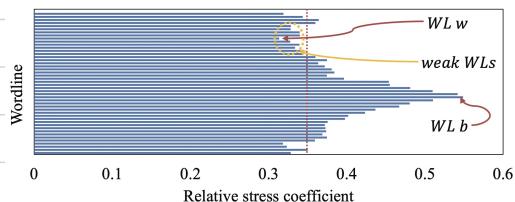


Figure 9: Per-WL relative stress coefficients.

- 그림 9에 용역( $S_k$ )가 클수록 더욱 강한 WL이 드.

## 4.3 Per-block Erase Modes

- low stress erase mode를 효과적으로 관리하기 위해 9개의 블록 삭제 모드를 제안. 표 3이 이게 대한 내용이다.

Table 3: A summary of nine gErase modes.

$n_m$  = Number of erase

$f$  = erase relief ratio

$K$  = block capacity reduction ratio

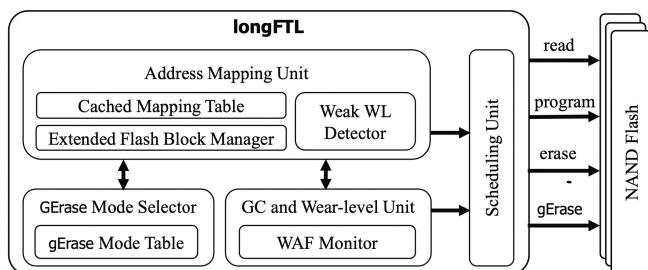
	gErase mode(n) == gE(n)								
	gE(1)	gE(2)	gE(3)	gE(4)	gE(5)	gE(6)	gE(7)	gE(8)	gE(9)
No. of protected WLs	8	12	16	20	24	24	24	28	32
Erase relief ratio $f$	25%	33%	38%	40%	42%	50%	50%	50%	50%
Block capacity reduction	1.04%	2.08%	3.13%	4.17%	5.21%	6.25%	7.29%	8.33%	9.38%
Norm. Max P/E cycles	1.19	1.26	1.30	1.33	1.37%	1.39	1.41	1.43	1.45

- gE(1) 모드로 지우기 되면 block의 수명은 45% 증가하여 gE(1)로 지울 경우 19%만 증가하게 된다.

• 높이 증가할 경우 수명이 증가하지만 블록의 용량은 감소하게 된다.

## 5. Design and Implementation of Long FTL

### 5.1 overview



• WL detector는 약한 WL을 탐색하는 역할을 하며  
WAF monitor는 WAF 변화를 감지한다. Extended Flash Block  
Manager에는 적절한 GErase 모드를 선택한다.

Figure 10: An organizational overview of longFTL.

### 5.2 Weak WL Detector

- N개의 WL 삭별을 요청 받았을 때 BER을 기준으로 WL을 선택하는 모듈

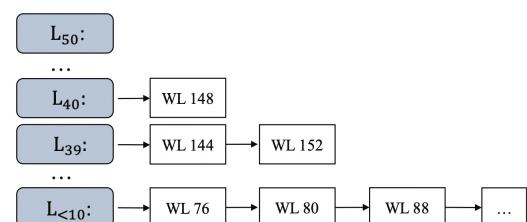


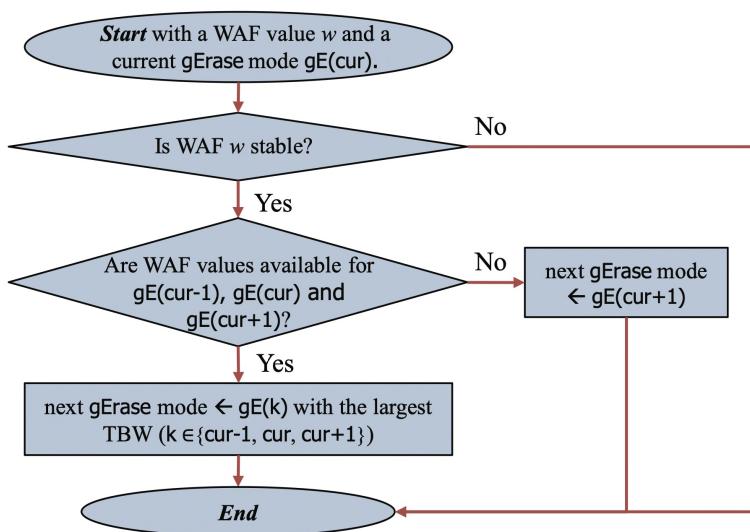
Figure 11: BER-sorted linked lists for detecting weak WLS.

- ECC 오류를 통해 약한 WL들을 linked-list 형태로 관리 한다.
- 1 bit-flag를 활용하여 해당 WL이 low-stress erase mode인지 파악한다.
- 이온 전류 용량의 의미한 부분만 차지하고 있다.
- 약한 WL은 100B/G 주제로 삭별하여 overhead를 최소화 하였다.
- 약한 WL을 찾기 위해 time overhead를 최소화
- Free한 block을 선택 할 때는 gErase가 가장 작은 Count를 선택

### 5.3 WAF monitor

- SSD의 WAF를 monitoring 하는 모듈
- SSD의 capacity가 감소하였다면 (공격적인 gErase에 의해) I/O 성능 감소에 영향을 끼치게 된다.
- 또한 예상대로 작동한다면 gErase 영향이 매우 감소하게 된다.
- I/O 성능 변동을 방지하기 위해서 I/O workload가 안정적인 상태일 때 gErase 모드를 수정하게 한다.

## 5.4 GErase Mode Selector



- 새로운 WAF가 축소된다면 그림 12의 절차를 흐름한다.
- WAF가 안정적이면 cur-1, cur, cur+1의 TBW를 측정하여 그에 TBW를 gErase 모드로 선택한다.
- TBW에 따라 +1 -1 유지를 결정한다.

Figure 12: A procedure for selecting the next gErase mode.

## 5.5 Extended Flash Block Manager

- Extended Flash Block Manager은 gErase를 평가하는 성능에 부정적인 영향을 주는지 점검한다.
- hot data 세팅 이용하여 Col-한 데이터에서는 용량 감소와 같은 부정적인 영향으로 이용하지 않는다.
- GLS 시뮬레이션은 gErase mode를 저장하지 않고

## 6 Experimental Results

### 6.1 Experimental Settings

- 실험 환경은 MQSim을 이용하여 Baseline과 BPM을 비교할 것이다.
- Baseline은 일반적인 page-level mapping FTL이며 BPM은 이미 bad-page management를 포함한 것이다.

Table 4: I/O characteristics of traces used for evaluations.

• workloads는 테이블 4와 나온 것을 사용 한다.

	proj0	prxy1	prxy0	src10	proj2	OLTP	filesrvr	varmail
Read:Write	6:94	64:36	5:95	46:54	86:14	70:30	40:60	40:60
Total writes (GB)	144	725	54	302	169	639	249	312
WAF (without gErase)	1.08	1.16	1.24	2.66	3.55	1.89	2.49	2.98

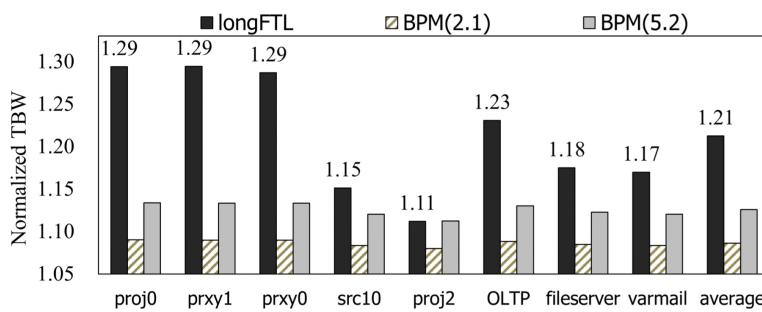
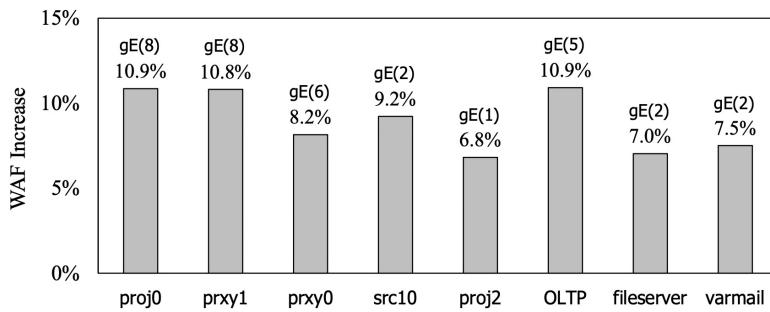


Figure 14: Comparisons of lifetime extensions.

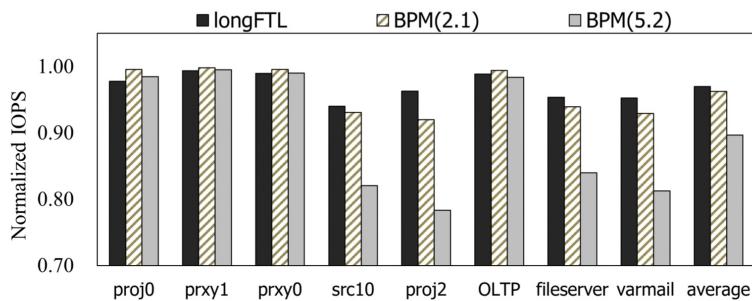
- workload에 따라 TBW가 다른 이유는 WAF가 다른 값을 가지고 있기 때문이다.
- longFTL은 단순한 BPM 보다 더욱 좋은 성능을 보여준다.
- WAF 영향에 따라 BPM과 유사한 성능을 보이는 결과도 존재한다.



- proj0, prxy1은 높은 WAF를 가지고 있어 대용량 기록을 보호할 것이다.
- Proj2는 WAF가 가장 작기 때문에 gE(1) 영향을 가장 미미하다.
- TBW에 따라 SVC10과 같은 결과가 나올 수도 있다.

Figure 15: Comparisons of WAF changes under dominant gErase modes used.

### 6.3 performance Overhead

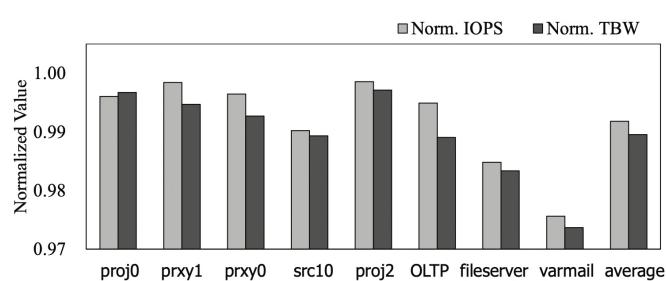


- long FTL은 높은 성능 저하를 보여준다.
- BPM은 수명 연장 효과가 뛰어나기 때문에 long FTL이 더욱 효과적인 결과를 보여줄 수 있다.

Figure 16: Comparisons of performance overhead.

Figure 17 shows how the more aggressive policy works over the current one. All the values were normalized over those of longFTL. The evaluation result shows that TBW and IOPS can degrade up to 3% (as in varmail). In the aggressive policy, the TBW improvement ratio can be reduced by up to 15.2%<sup>13</sup> over the current conservative policy.

### 6.4 Effectiveness of Block Relief policy



- 일반적인 GEA G(S) G(E)를 통해서 적용을 비교
- TBW와 IOPS의 성능 저하를 비교한다.
- TBW는 proj 15.2% 감소할 수 있다.

→ 따라서 Hot 데이터만 적용하는 것이 좋지?

Figure 17: Performance and lifetime comparisons of an aggressive gErase application policy.

## 1 Related Work

- SSB는 endurance를 높이는 방식으로 빠른 연수가 이루어졌다.
- 지속 전망을 갖게 하여 수명을 증대시킨 연구가 있다. 하지만 세분화된 특징을 이용하고 있지 않아.
- 프로그램 자체를 집중한 연구가 존재하지만 마모의 주된 원인은 단계적이며 향상되거나  
않아.
- BIM 대신 BFM 기반으로 계획되었지만 이는 큰 성능 저하를 포함할 수 있다.

## 2 Conclusion

- WL의 내구성을 활용하여 성능 저하를 최소화하는 플래시 면장 기술인 GEM을 제안
- Long ITL에 대한 디자인 윈드를 개발하여 21% 수명 연장 효과를 나타낸다.
- 미래에는 기계 효율 기반 GEM을 도입 하겠다.