Write Amplification due to ECC on Flash Memory or Leave those Bit Errors Alone

Sangwhan Moon and A. L. Narasimha Reddy

Texas A&M University

sangwhan@tamu.edu **reddy@ece.tamu.edu**

INTRODUCTION (1/2)

- Flash Memory Write Endurance Problem
  - 10,000 P/E cycles for MLC

- Flash Memory Protection Scheme
  - Error Correcting Code (ECC)
  - Scrubbing
  - Wear-leveling and Garbage Collection

- These protection schemes

  (+) Improve the reliability of flash memory

  (–) Amplify writes → Reduce the reliability of flash memory

INTRODUCTION (2/2)

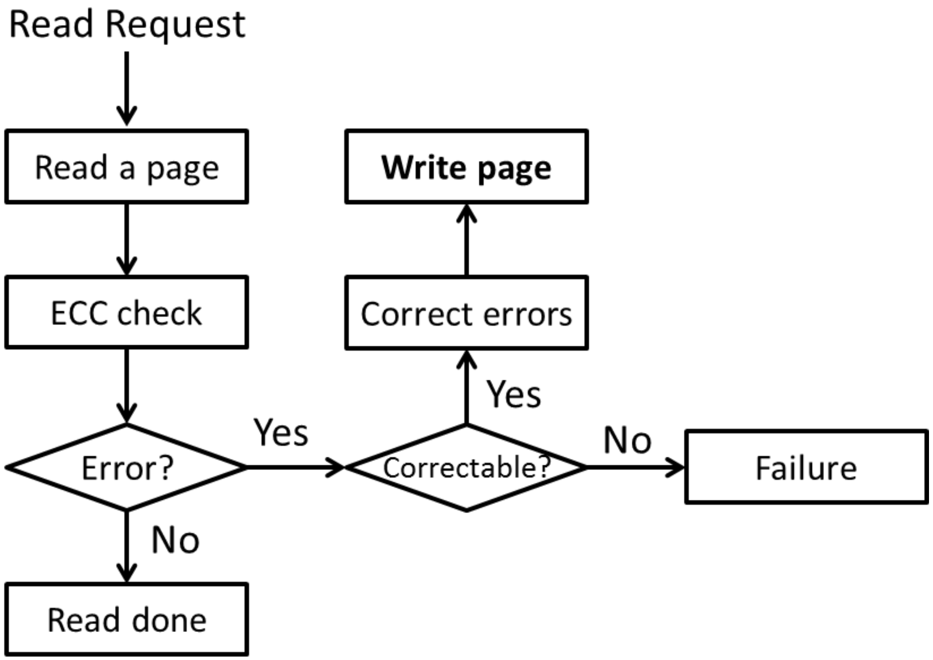
- **Write amplification**
  - Writes internally done / Writes externally issued

- Main sources
  - Copying live data in *garbage collection* (prior work)
  - Writing corrected data back in ***ECC recovery***

- Write amplification degrades
  - *write performance* (prior work)
  - flash memory's ***lifetime***

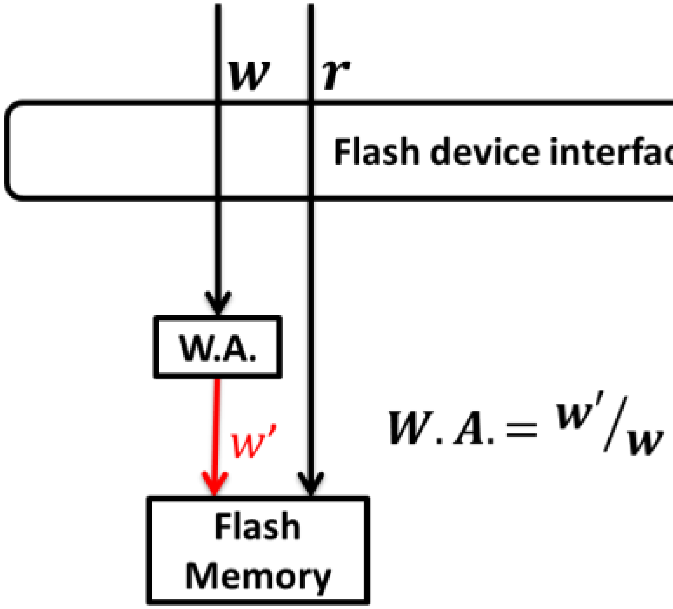
WRITE AMPLIFICATION FROM ECC

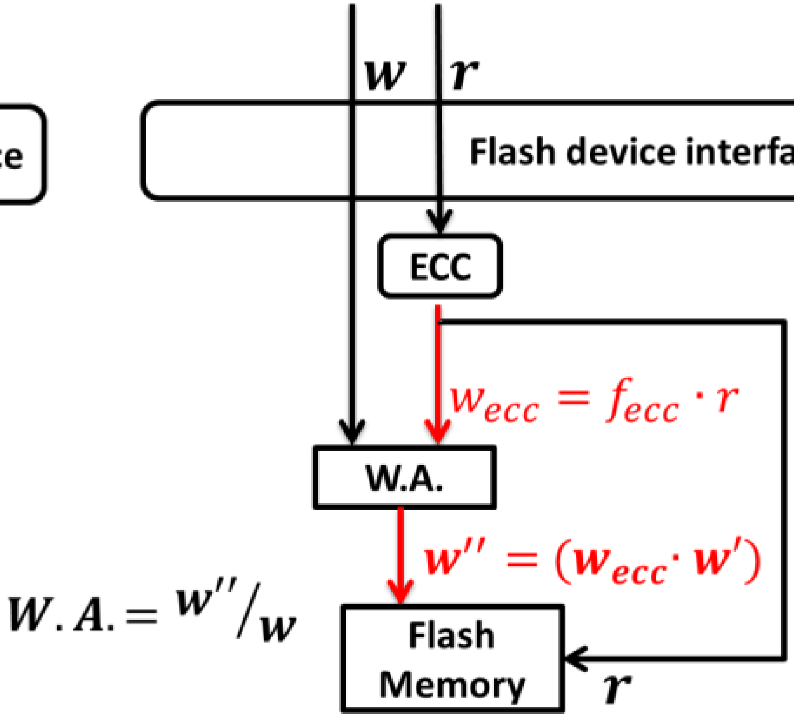
- W.A. due to ECC recovery
  - Reads lead to writes



WRITE AMPLIFICATION FROM ECC

- A traditional point of view to WA and our point of view to WA
- Severe problem with read intensive workload



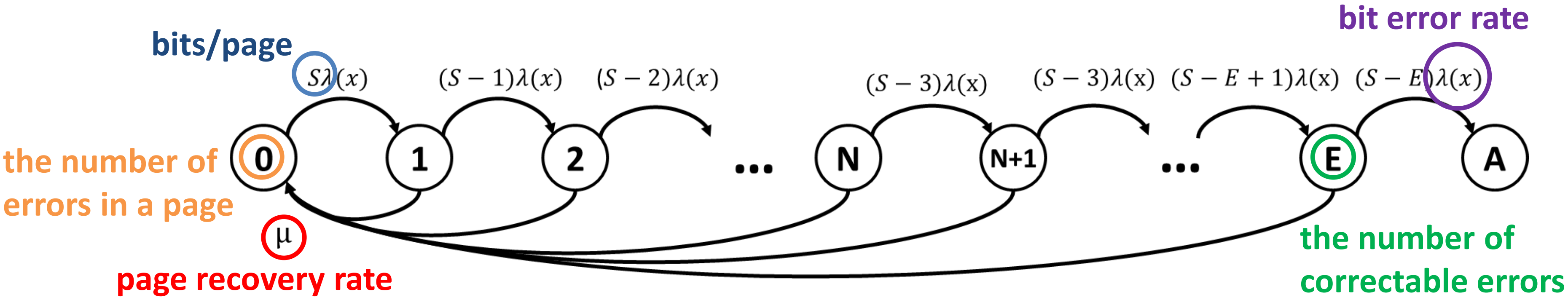


CONTRIBUTION

- A statistical model
  - The impact of the W.A. to the lifetime of flash
- A loss of 50% of the lifetime due to the W.A.
  - 20% due to garbage collection, 30% due to ECC
- Threshold-based ECC to reduce the W.A.
  - Improves the lifetime up to 40%.

A Reliability Model

- Raw Bit Error Rate from measurement study
- A Canonical Markov Model


$S\lambda(x)$
$(S-1)\lambda(x)$
$(S-2)\lambda(x)$
$(S-3)\lambda(\mathrm{x})$
$(S-3)\lambda(\mathrm{x})$
$(S-E+1)\lambda(\mathrm{x})$
$(S-E)\lambda(x)$
0
1
2
...
N
N+1
...
E
A
μ

bits/page

bit error rate

the number of errors in a page

page recovery rate

the number of correctable errors

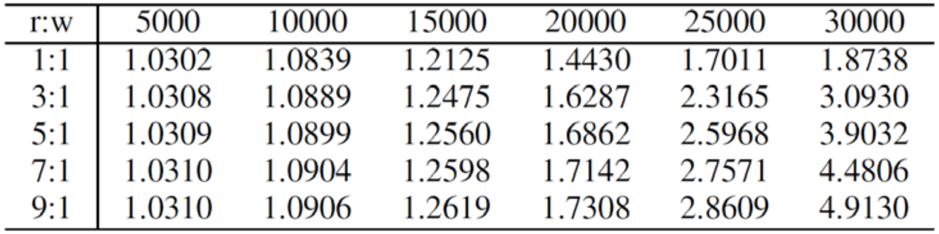
- Mean Time To Data Loss

$$\mathrm{MTTDL_p} = \lim_{k\to\infty} \sum_{j=1}^{k} \left(\mathrm{j g(j)} \prod_{i=1}^{j-1} (1 - \mathrm{g(i)}) \right)$$

The probability of getting into the absorbing state A in the Markov chain

EVALUATION

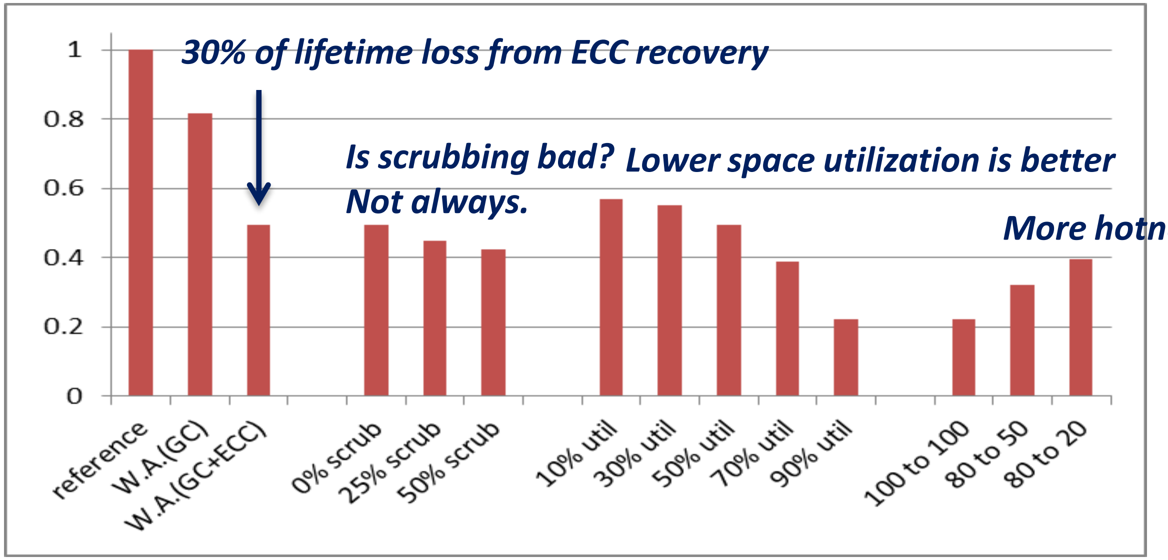
- WA from ECC recovery
- Scrubbing
- Space utilization
- Hot/cold dichotomy



r:w	5000	10000	15000	20000	25000	30000
1:1	1.0302	1.0839	1.2125	1.4430	1.7011	1.8738
3:1	1.0308	1.0889	1.2475	1.6287	2.3165	3.0930
5:1	1.0309	1.0899	1.2560	1.6862	2.5968	3.9032
7:1	1.0310	1.0904	1.2598	1.7142	2.7571	4.4806
9:1	1.0310	1.0906	1.2619	1.7308	2.8609	4.9130

W.A. from ECC recovery at different P/E cycles

160GB 3x nm SSD
100MB/s Bandwidth
61bits correctable / 4KB
50% Random Workload
50% Device Utilization
R:W=3:1



30% of lifetime loss from ECC recovery

Is scrubbing bad? Not always.

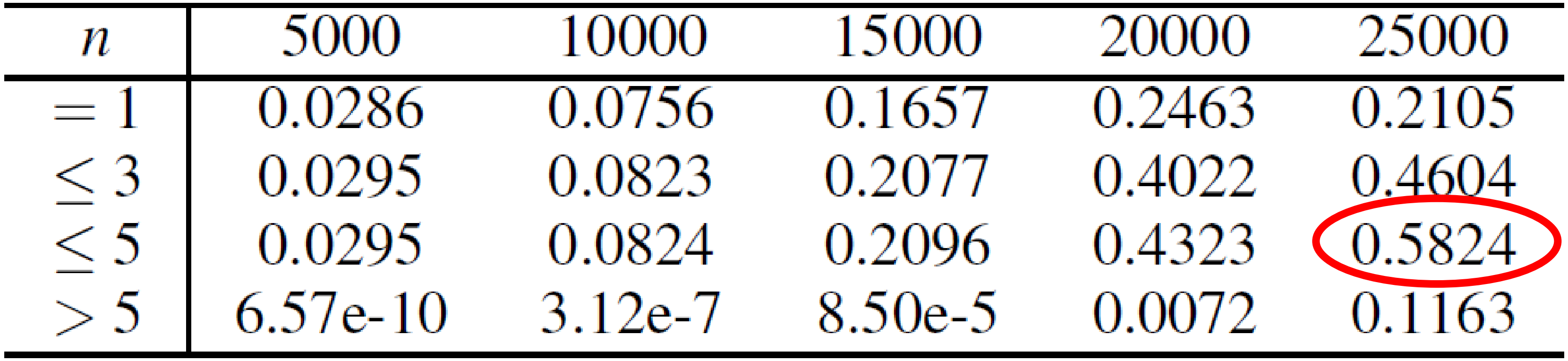
Lower space utilization is better

More hotness is better

Threshold-based ECC (1/3)

- A few bit errors accumulate before ECC correction

58.2% of recoveries
for pages with <= 5 bit errors



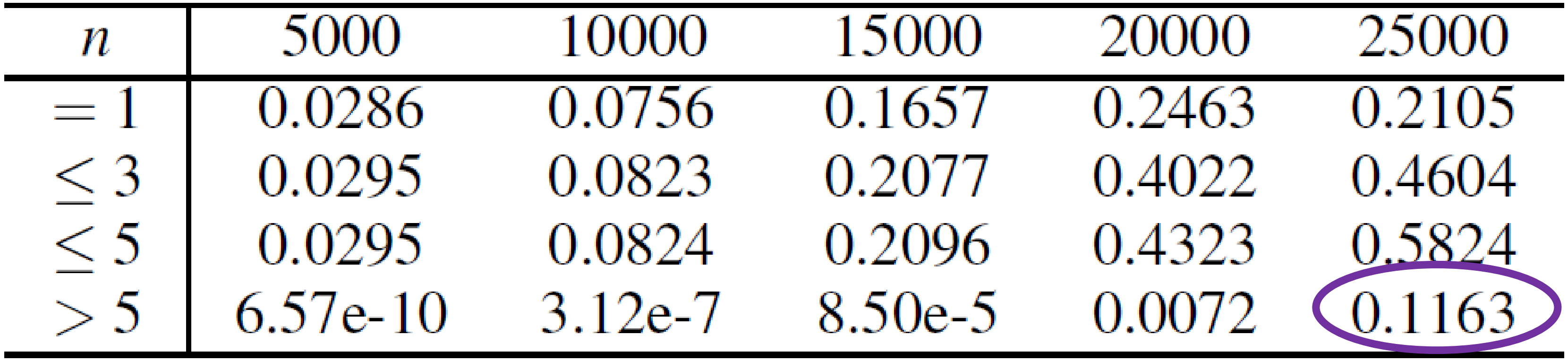
n	5000	10000	15000	20000	25000
$= 1$	0.0286	0.0756	0.1657	0.2463	0.2105
≤ 3	0.0295	0.0823	0.2077	0.4022	0.4604
≤ 5	0.0295	0.0824	0.2096	0.4323	0.5824
> 5	6.57e-10	3.12e-7	8.50e-5	0.0072	0.1163

Probability distribution of the number of accumulated bit errors *n* when they are recovered by ECC

Threshold-based ECC (1/3)

- A few bit errors accumulate before ECC correction

11.6% of recoveries
for pages with > 5 bit errors

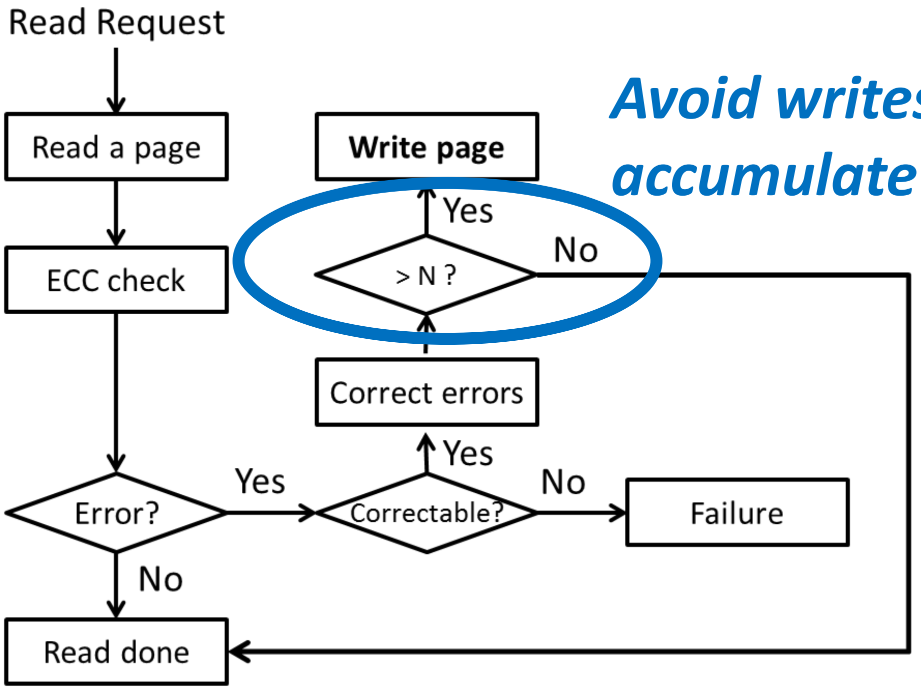


n	5000	10000	15000	20000	25000
$= 1$	0.0286	0.0756	0.1657	0.2463	0.2105
≤ 3	0.0295	0.0823	0.2077	0.4022	0.4604
≤ 5	0.0295	0.0824	0.2096	0.4323	0.5824
> 5	6.57e-10	3.12e-7	8.50e-5	0.0072	0.1163

Probability distribution of the number of accumulated bit errors *n* when they are recovered by ECC

THRESHOLD-BASED ECC (2/3)

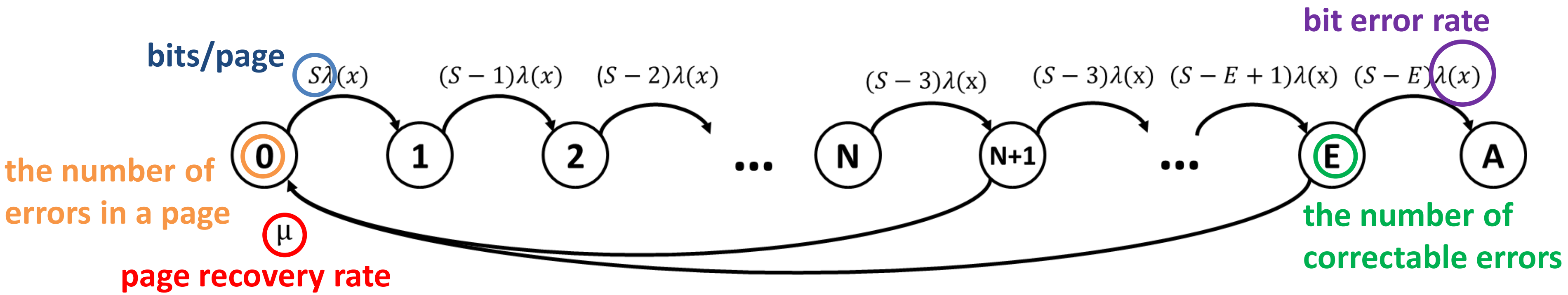
- Postpone write until errors accumulate?



Avoid writes until bit errors accumulate to a threshold

Threshold-based ECC (3/3)

- Reliability Model



bits/page

bit error rate

the number of errors in a page

page recovery rate

the number of correctable errors

- Evaluation

Optimal Threshold

Threshold(%)	0	10	30	50	70	90
R.MTTDL	0.496	0.614	0.671	0.694	0.702	0.696

CONCLUSION

- Reads lead to the W.A.
  - A Statistical Reliability Model
  - A loss of 30% of the lifetime due to ECC recovery under 50% workload and R:W = 3:1.
- To control the W.A. through two tools
  - *Scrubbing* for detecting latent errors
  - *Threshold-based ECC* for avoiding excessive recovery

Thank you!

Questions and Answers?

TEXAS A&M UNIVERSITY