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 at tamu.edu
reddy at ece.tamu.edu

Introduction (1/2)

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

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

- These protection schemes

  (+) Improve the reliability of flash memory

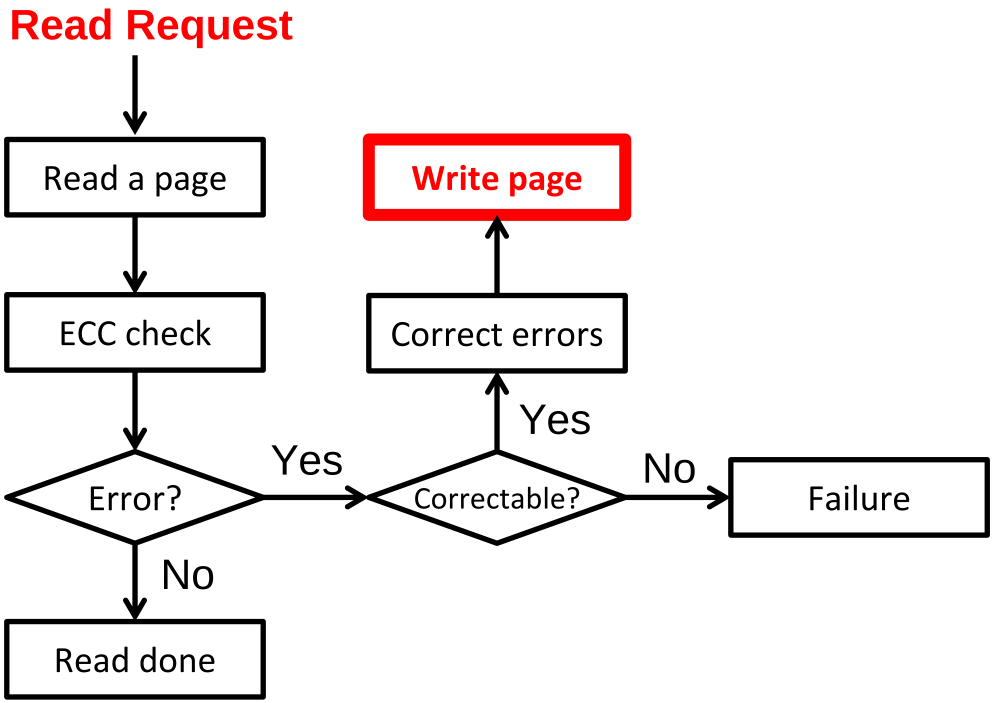
  (–) Amplify writes → Reduce the reliability of flash memory

Introduction (2/2)

- **Write amplification (W.A.)**
  - The number of excess writes / writes issued by system
- Main sources of W.A.
  - Copying live data in **garbage collection**
  - Writing corrected data back in **ECC recovery**
  - Parity update of **RAID**
- W.A. degrades
  - performance (related work)
  - flash memory’s **lifetime**

Write Amplification from ECC (1/2)

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



Read Request

Read a page

ECC check

Error?

No

Read done

Yes

Correctable?

Yes

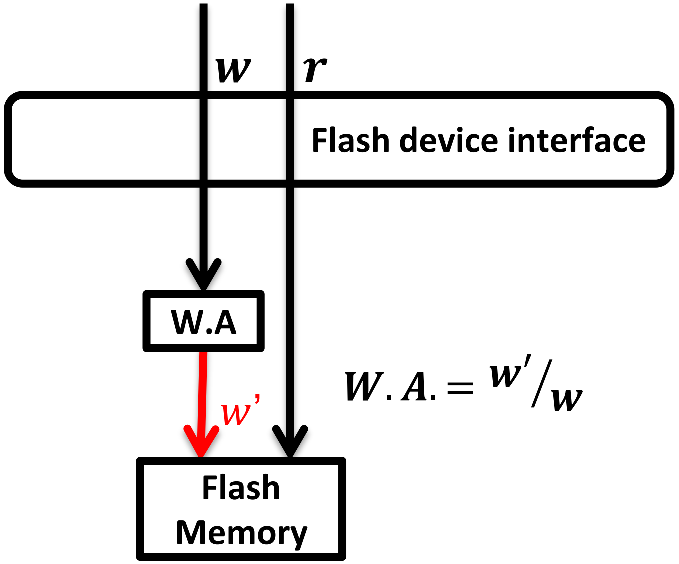
Correct errors

Write page

No

Failure

Write Amplification from ECC (2/2)



w

r

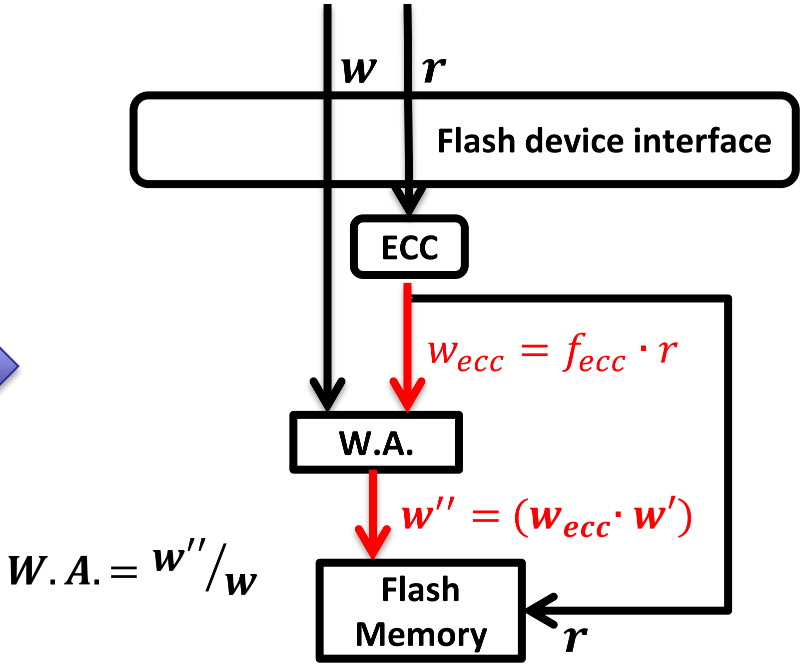
Flash device interface

W.A

w'

Flash Memory

$$W.A. = {w'}/{w}$$



w

r

Flash device interface

ECC

$w_{ecc} = f_{ecc} \cdot r$

W.A.

$w'' = (w_{ecc} \cdot w')$

Flash Memory

r

$$W.A. = {w''}/{w}$$

Traditional point of view to W.A.

Our point of view to W.A.

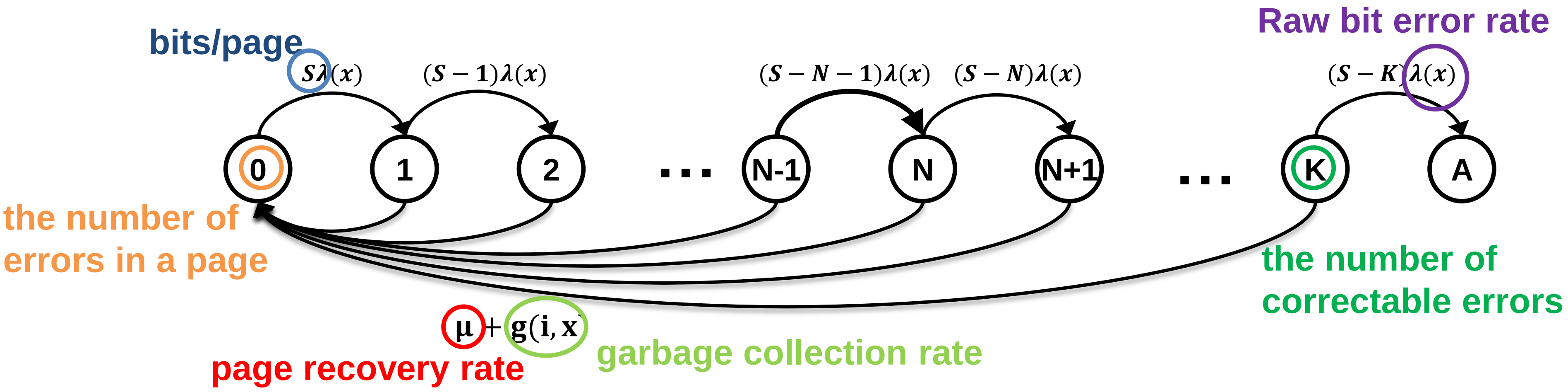
→ ***Severe problem with read intensive workload***

Contribution

- A statistical model
  - The impact of the W.A. to the lifetime of flash memory
- A loss of 48% of the lifetime due to the W.A.
  - Various parameters were tested
- Threshold-based ECC to reduce the W.A.
  - Improves the lifetime up to 64%
  - The way to control W.A. to maximize the lifetime

A Reliability Model (1/2)

- Raw Bit Error Rate from measurement study
- Uncorrectable Page Error Rate (UPER)
  - A Canonical Markov Model



bits/page

Raw bit error rate

the number of errors in a page

the number of correctable errors

page recovery rate

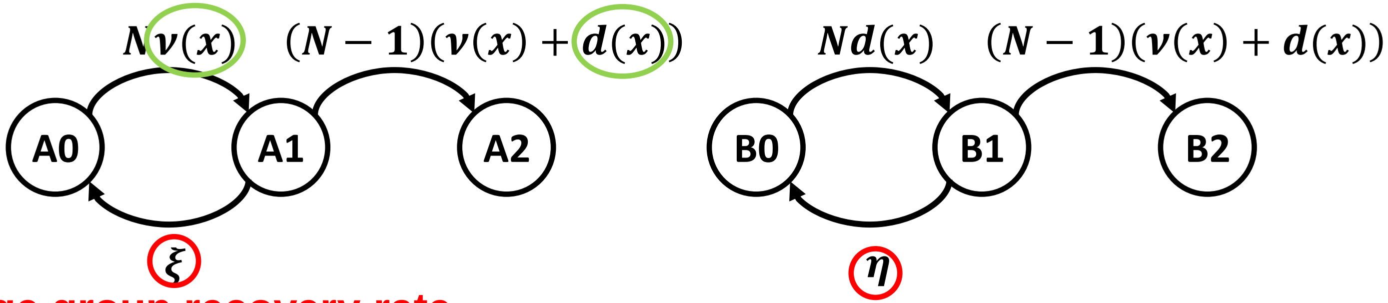
garbage collection rate

A Reliability Model (2/2)

- Uncorrectable Page Group Error Rate (UGER)

UPER

whole device failure rate



page group recovery rate

device recovery rate

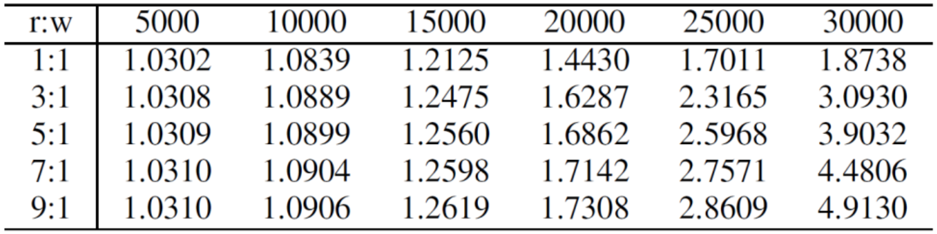
- Mean Time to Data Loss

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

The probability of
uncorrectable page group error

Evaluation: Write Amplification

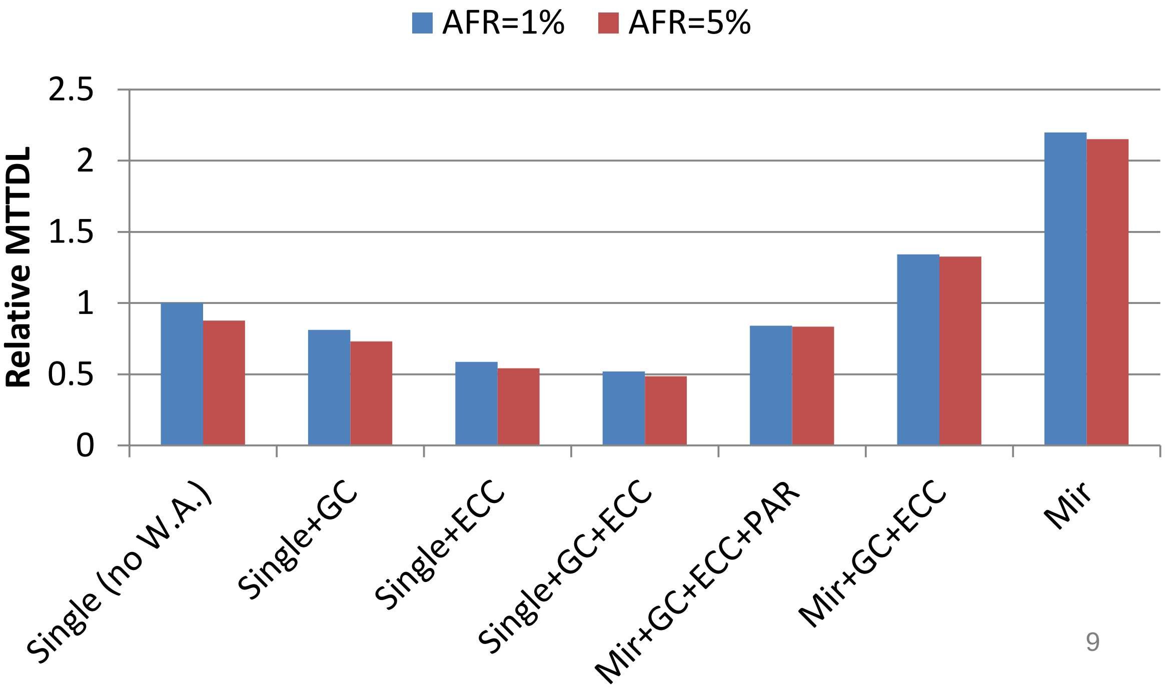
- More read, higher W.A.
- G.C. : -19% lifetime
- ECC : -42% lifetime
- G.C. + ECC : -48% lifetime



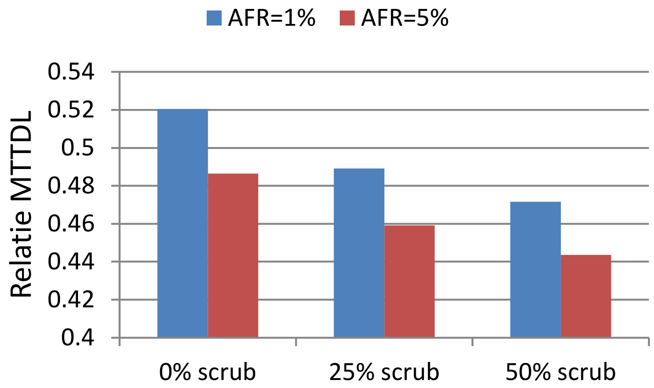
r:w	5000	10000	15000	20000	25000	30000
1:1	1.0302	1.0839	1.2125	1.4430	1.7011	1.8738
3:1	1.0308	1.0889	1.2475	1.6287	2.3165	3.0930
5:1	1.0309	1.0899	1.2560	1.6862	2.5968	3.9032
7:1	1.0310	1.0904	1.2598	1.7142	2.7571	4.4806
9:1	1.0310	1.0906	1.2619	1.7308	2.8609	4.9130

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

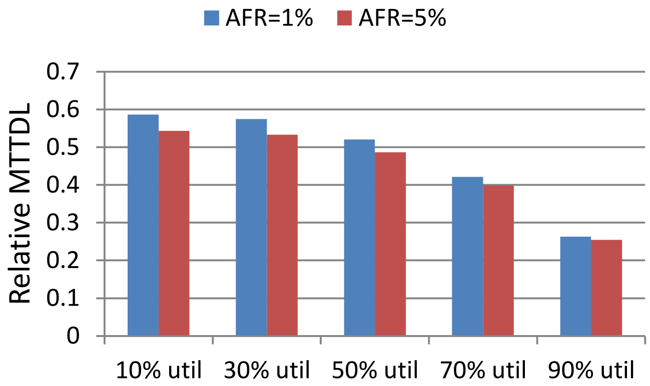
160GB 3x nm SSD
100MB/s Bandwidth
61bits correctable / 4KB
50% Device Utilization
R:W=3:1



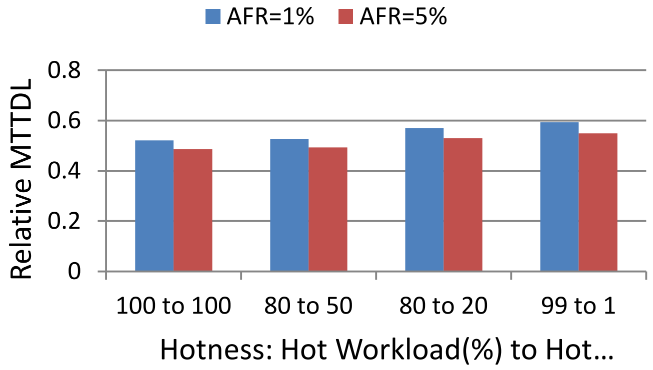
Evaluation: Various Parameters



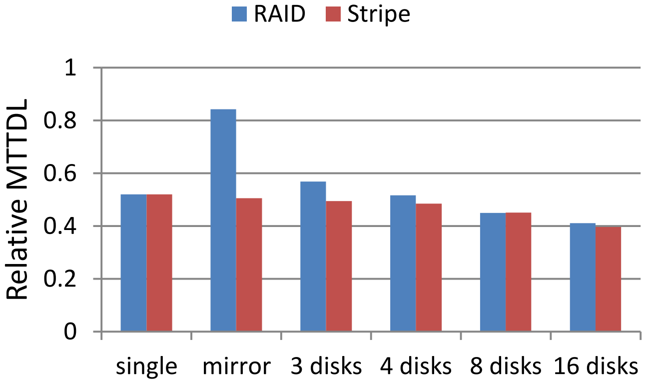
1. Scrubbing may be harmful (not always)



2. Lifetime highly depends on space utilization



3. Hotness helps to improve lifetime
(efficient garbage collection)

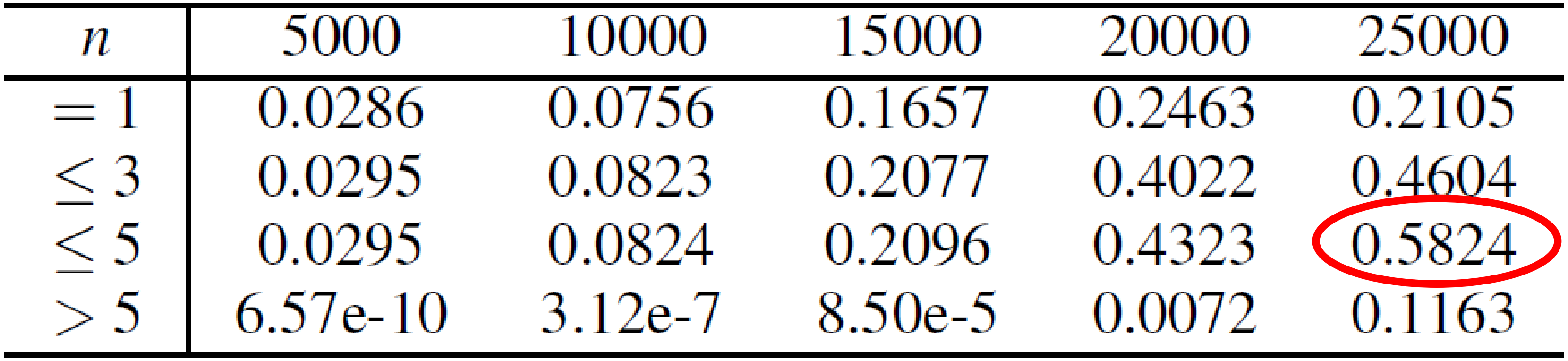


4. RAID improves lifetime
(Mirroring is the best since it splits read workload)

Threshold-based ECC (1/5)

- 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 (1/5)

- A few bit errors accumulate before ECC correction

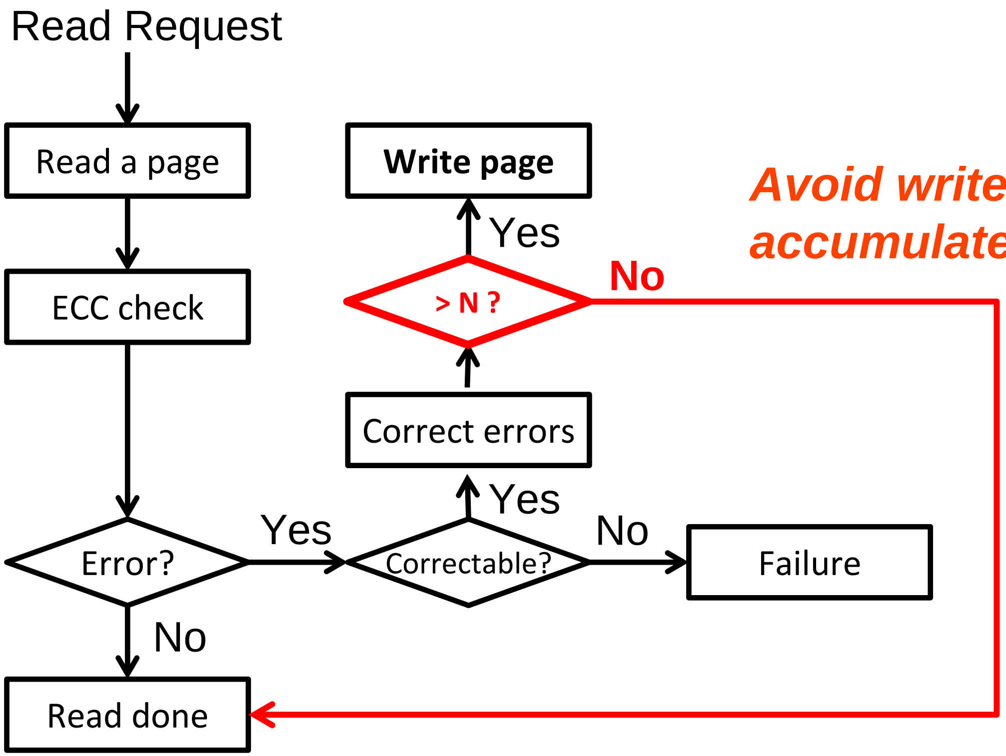
11.6% of recoveries
for pages with > 5 bit errors

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 (2/5)

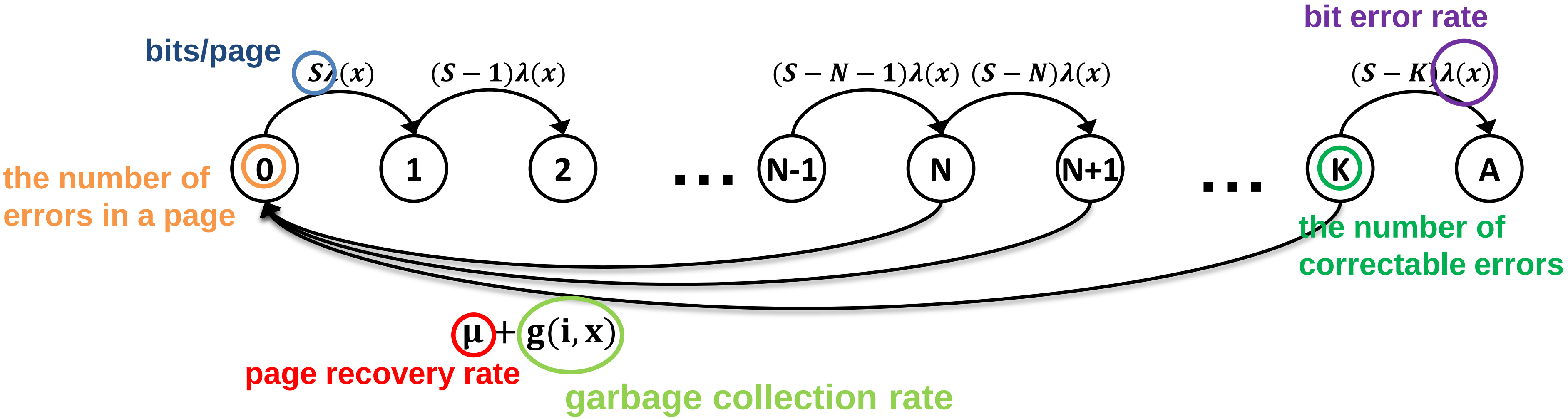
- IDEA: Postpone write until errors accumulate?



Avoid writes until bit errors accumulate to a threshold

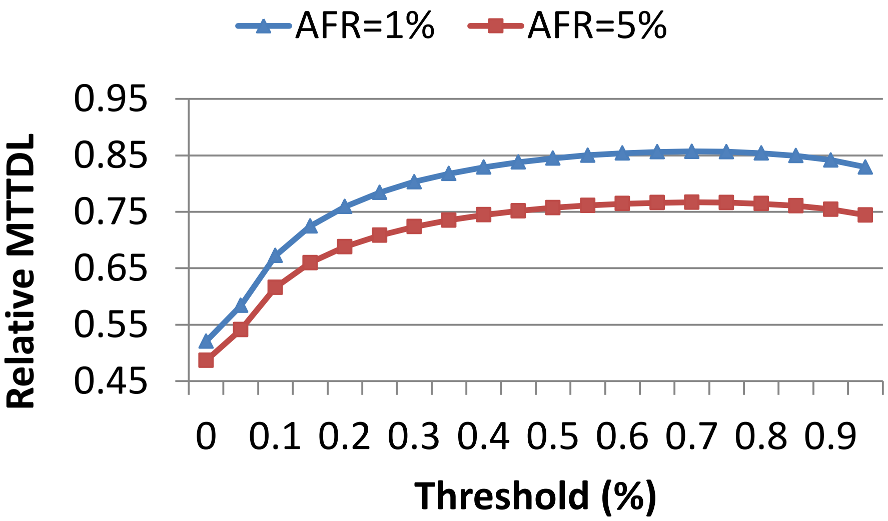
Threshold-based ECC (3/5)

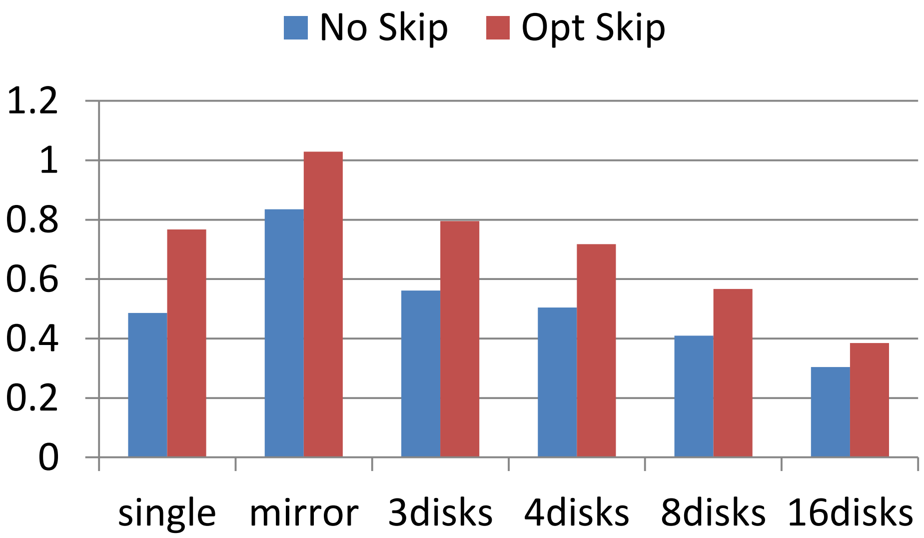
- A Markov model for reliability analysis



Threshold-based ECC (4/5)

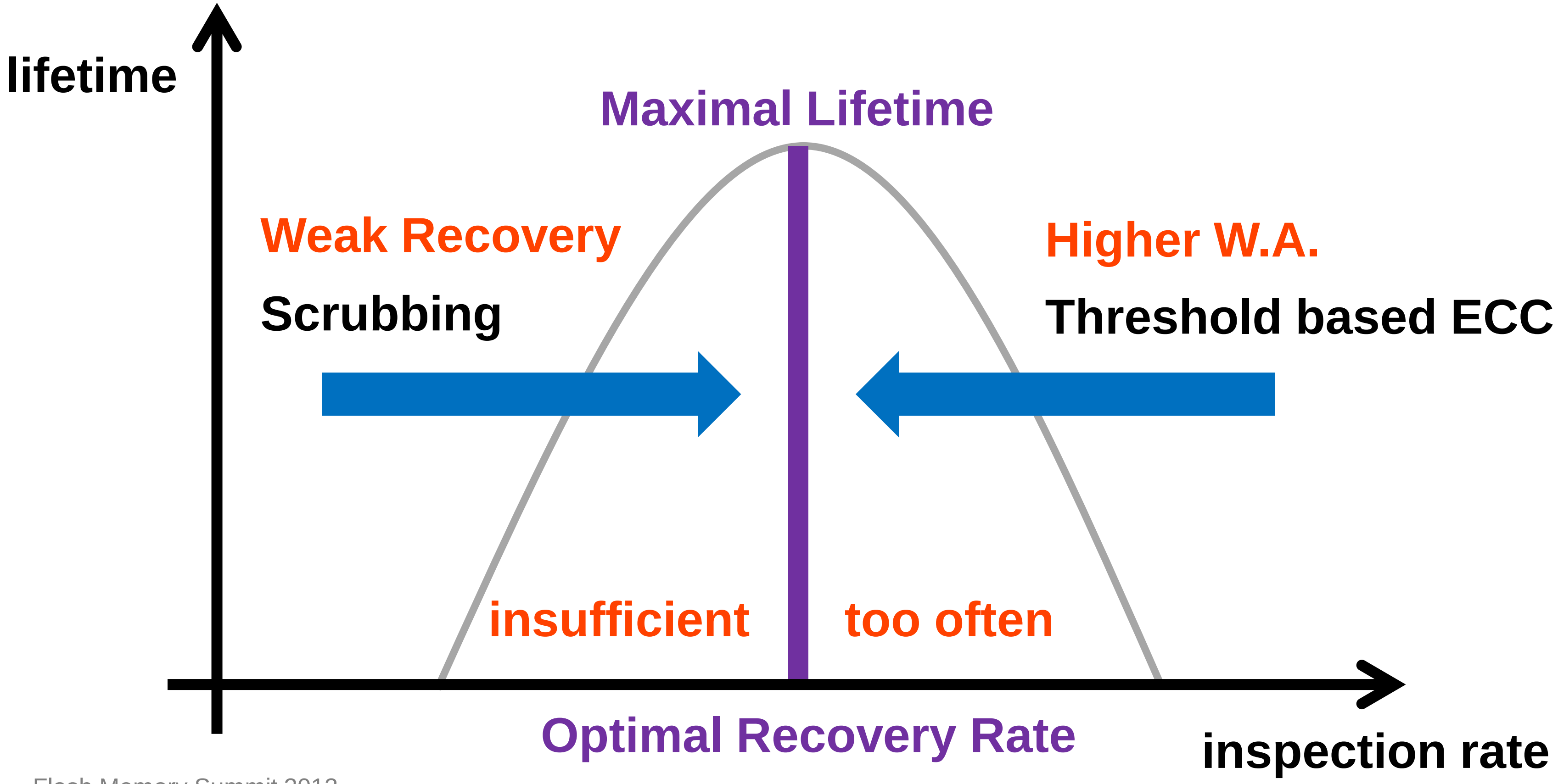
- Evaluation
  - Optimal threshold depends on environment and # of devices
  - Lifetime improves up to 64%





Threshold-based ECC (5/5)

- Controlling W.A. to achieve max lifetime



Conclusion

- Reads lead to W.A.
  - A Statistical Reliability Model
  - A loss of 48% of the lifetime due to W.A.

- To control W.A. through two tools
  - Scrubbing for detecting latent errors
  - **Threshold-based ECC** for avoiding excessive recovery

Thank you

Q & A

Sangwhan Moon (sangwhan at tamu.edu)
A. L. Narasimha Reddy (reddy at ece.tamu.edu)