Computer Architecture

Lecture 15a: Flash Memory and Solid-State Drives (II)

Prof. Onur Mutlu
ETH Zürich

Fall 2018

14 November 2018

Readings on Flash Memory

More Background and State-of-the-Art

INVITED PAPER

Proceedings of the IEEE, Sept. 2017

Error Characterization, Mitigation, and Recovery in Flash-Memory-Based Solid-State Drives

This paper reviews the most recent advances in solid-state drive (SSD) error characterization, mitigation, and data recovery techniques to improve both SSD's reliability and lifetime.

By Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu

https://arxiv.org/pdf/1706.08642

More Up-to-date Version

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu,
"Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery"
Invited Book Chapter in Inside Solid State Drives, 2018.
[Preliminary arxiv.org version]

Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery

YU CAI, SAUGATA GHOSE
Carnegie Mellon University

ERICH F. HARATSCH
Seagate Technology

YIXIN LUO
Carnegie Mellon University

ONUR MUTLU
ETH Zürich and Carnegie Mellon University

Flash Memory Reliability

Agenda

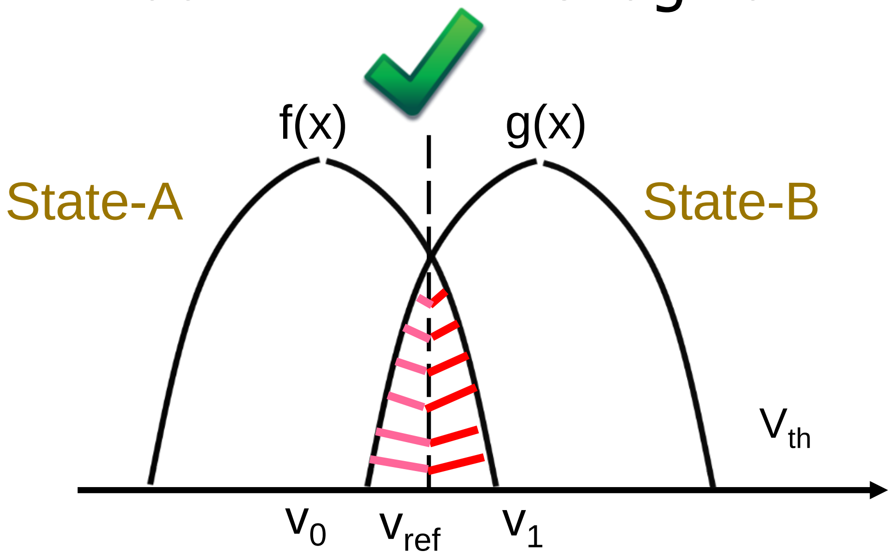
- Background, Motivation and Approach
- Experimental Characterization Methodology
- Error Analysis and Management
  - Main Characterization Results
  - Retention-Aware Error Management
  - Threshold Voltage and Program Interference Analysis
  - Read Reference Voltage Prediction
  - Neighbor-Assisted Error Correction
  - Read Disturb Error Handling
  - Retention Error Handling
  - 3D NAND Flash Memory Reliability
- Summary

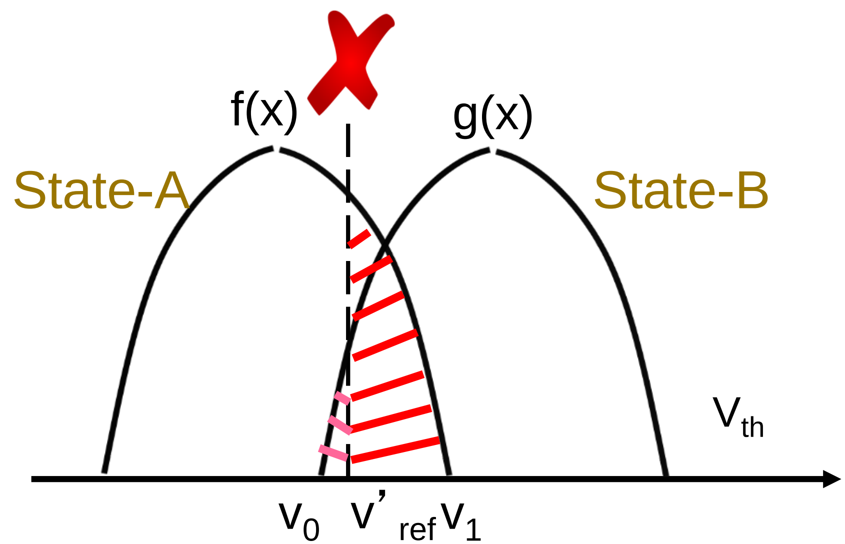
Using the Vth Distribution Models

- So, what can we do with the model?
- Goal: Mitigate the effects of program interference caused voltage shifts

Optimum Read Reference for Flash Memory

- Read reference voltage affects the raw bit error rate



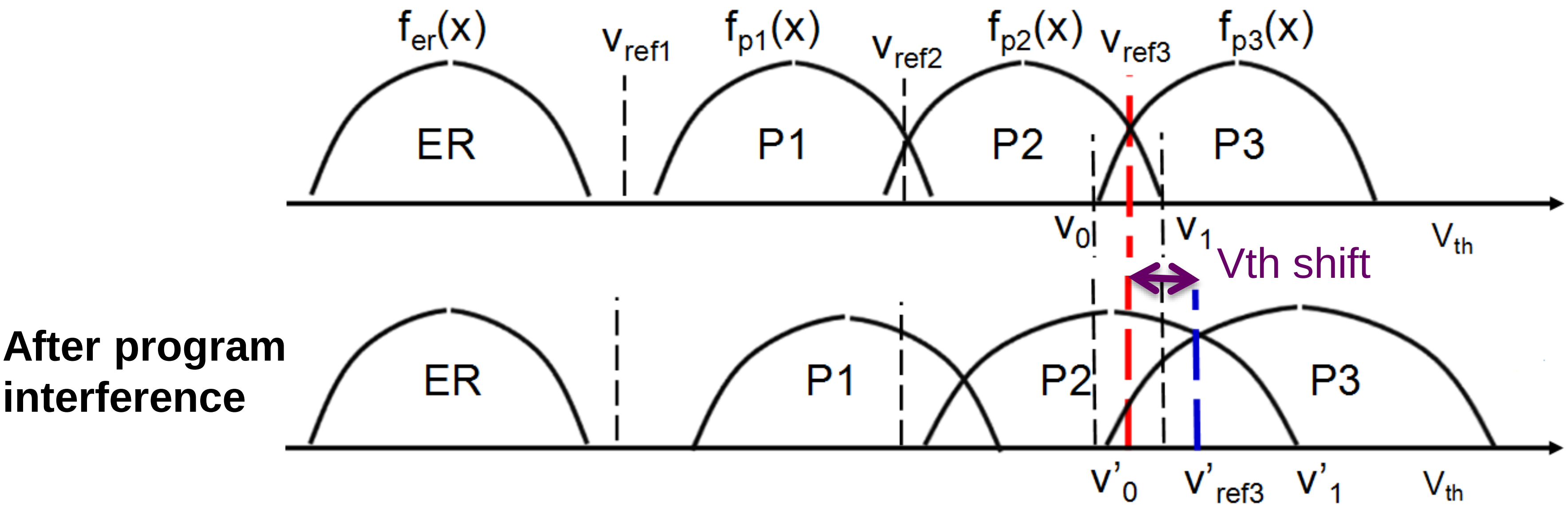


$$BER1 = \int_{v_{ref}}^{+\infty} f(x)dx + \int_{-\infty}^{v_{ref}} g(x)dx$$

$$BER2 = \int_{v'_{ref}}^{+\infty} f(x)dx + \int_{-\infty}^{v'_{ref}} g(x)dx$$

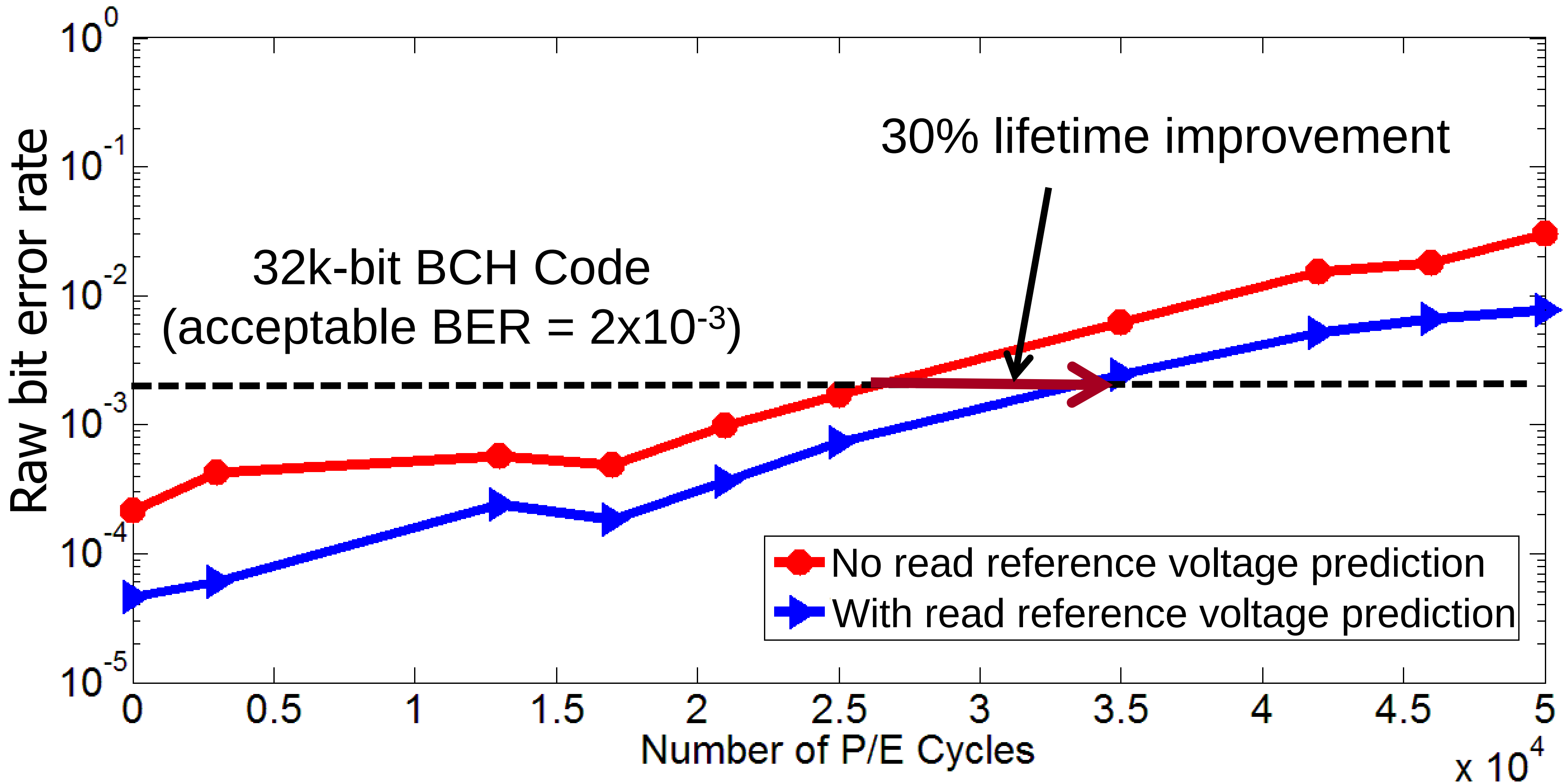
- There exists an optimal read reference voltage
  - Predictable if the statistics (i.e. mean, variance) of threshold voltage distributions are characterized and modeled

Optimum Read Reference Voltage Prediction



- Vth shift learning (done every ~1k P/E cycles)
  - Program sample cells with known data pattern and test Vth
  - Program aggressor neighbor cells and test victim Vth after interference
  - Characterize the mean shift in Vth (i.e., program interference noise)
- Optimum read reference voltage prediction
  - Default read reference voltage + Predicted mean Vth shift by model

Effect of Read Reference Voltage Prediction



- Read reference voltage prediction reduces raw BER (by 64%) and increases the P/E cycle lifetime (by 30%)

More on Read Reference Voltage Prediction

- Yu Cai, Onur Mutlu, Erich F. Haratsch, and Ken Mai, **"Program Interference in MLC NAND Flash Memory: Characterization, Modeling, and Mitigation"** *Proceedings of the 31st IEEE International Conference on Computer Design* (***ICCD***), Asheville, NC, October 2013. Slides (pptx) (pdf) Lightning Session Slides (pdf)

Program Interference in MLC NAND Flash Memory: Characterization, Modeling, and Mitigation

Yu Cai[1], Onur Mutlu[1], Erich F. Haratsch[2] and Ken Mai[1]
1. Data Storage Systems Center, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA
2. LSI Corporation, San Jose, CA
yucaicai@gmail.com, {omutlu, kenmai}@andrew.cmu.edu

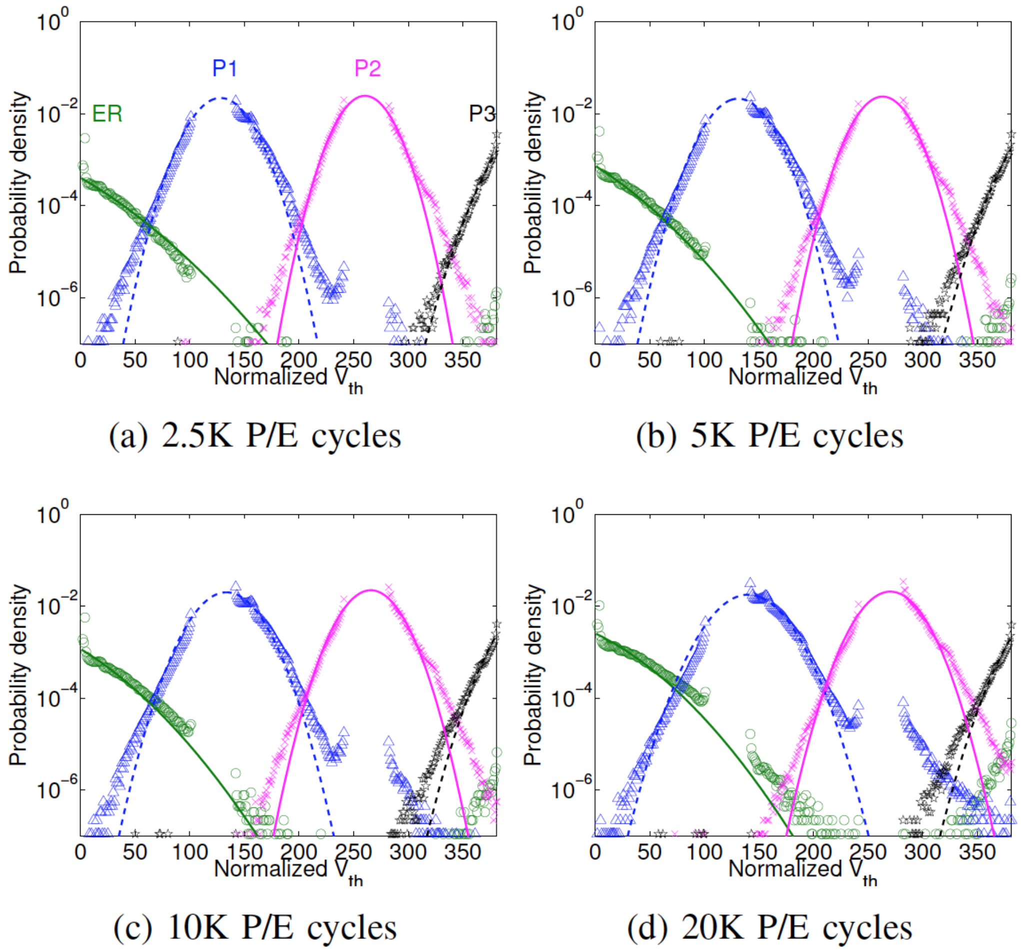
More Accurate and Online Channel Modeling

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu, **"Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory"** *to appear in IEEE Journal on Selected Areas in Communications (**JSAC**),* 2016.

Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory

Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, Onur Mutlu

Non-Gaussian Vth Distributions (1X-nm)



(a) 2.5K P/E cycles

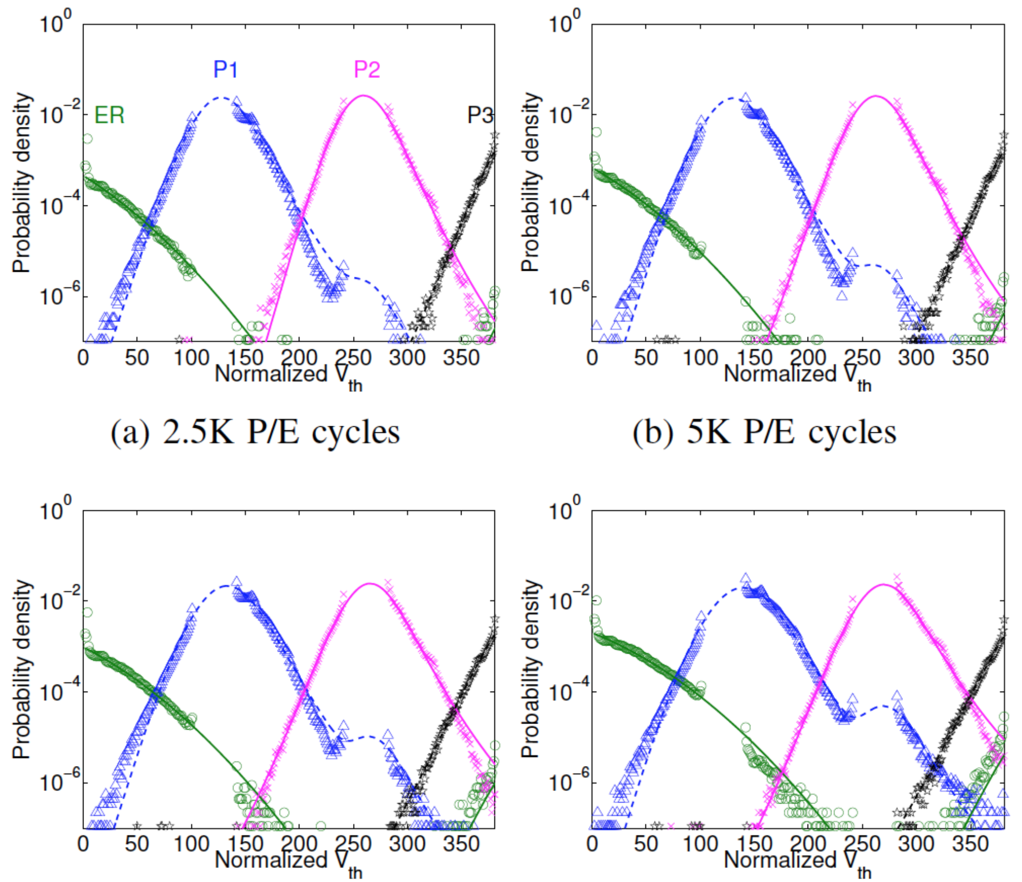
(b) 5K P/E cycles

(c) 10K P/E cycles

(d) 20K P/E cycles

Fig. 4: Gaussian-based model (solid/dashed lines) vs. data measured from real NAND flash chips (markers) under different P/E cycle counts.

Better Modeling of Vth Distributions (I)



(a) 2.5K P/E cycles
(b) 5K P/E cycles
(c) 10K P/E cycles
(d) 20K P/E cycles

Fig. 6: Our new Student's t-based model (solid/dashed lines) vs. data measured from real NAND flash chips (markers) under different P/E cycle counts.

Luo+, "Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory", JSAC 2016.

Better Modeling of Vth Distributions (II)

P/E Cycles	0	2.5K	5K	7.5K	10K	12K	14K	16K	18K	20K	AVG
Gaussian	.99%	1.8%	1.6%	1.8%	1.9%	2.4%	3.1%	8.7%	2.1%	2.3%	**2.6%**
Normal-Laplace	.34%	.46%	.55%	.61%	.63%	.67%	.68%	.70%	.67%	.67%	**.61%**
Student's t	.37%	.51%	.61%	.68%	.70%	.76%	.76%	.78%	.76%	.78%	**.68%**

TABLE 1: Modeling error of the evaluated threshold voltage distribution models, at various P/E cycle counts.

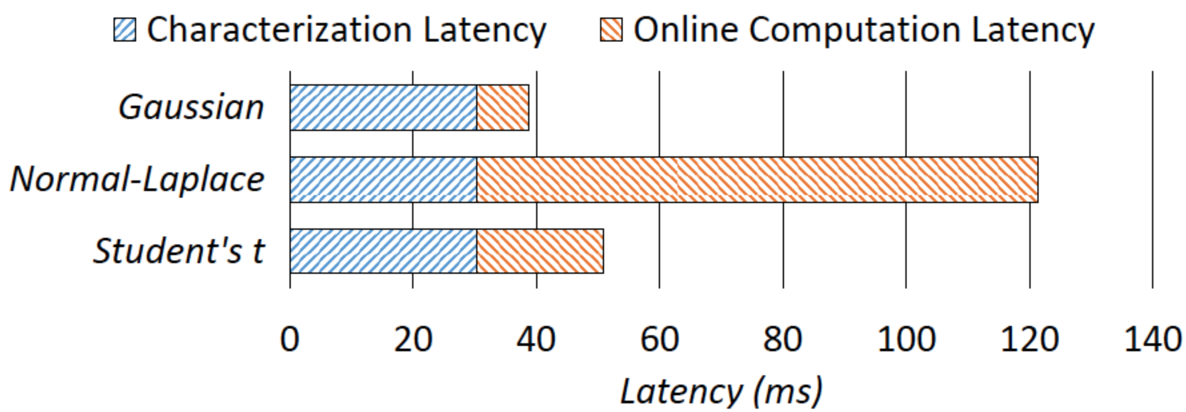


Fig. 8: Overall latency breakdown of the three evaluated threshold voltage distribution models for static modeling.

Vth Prediction vs. Reality with Better Modeling

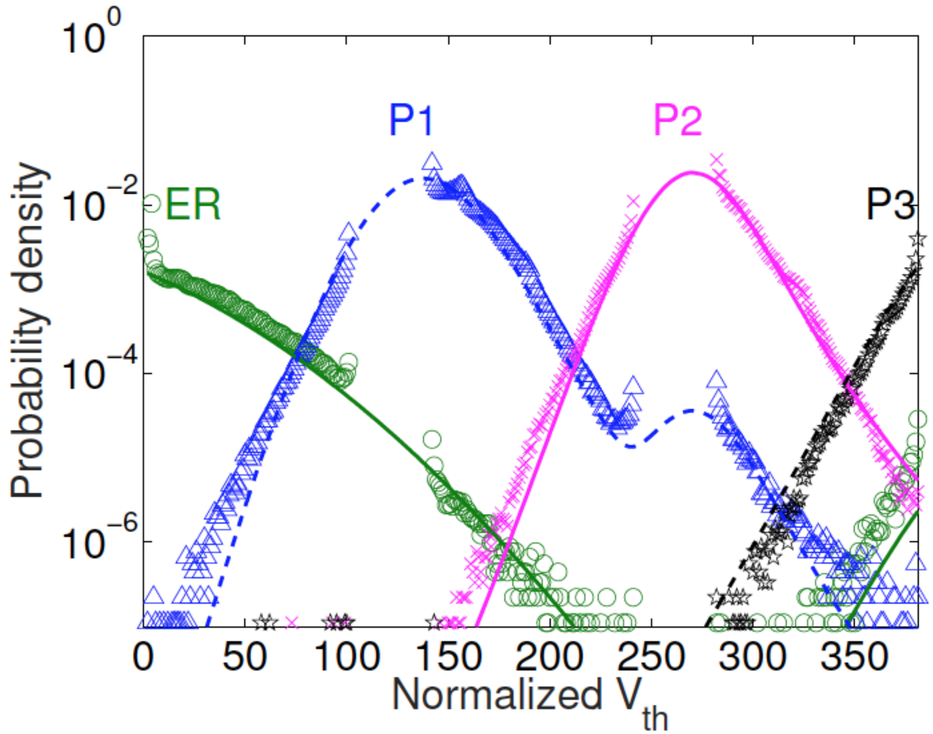


Fig. 13: Threshold voltage distribution as predicted by our dynamic model for 20K P/E cycles, using characterization data from 2.5K, 5K, 7.5K, and 10K P/E cycles, shown as solid/dashed lines. Markers represent data measured from real NAND flash chips at 20K P/E cycles.

Online Read Reference Voltage Prediction

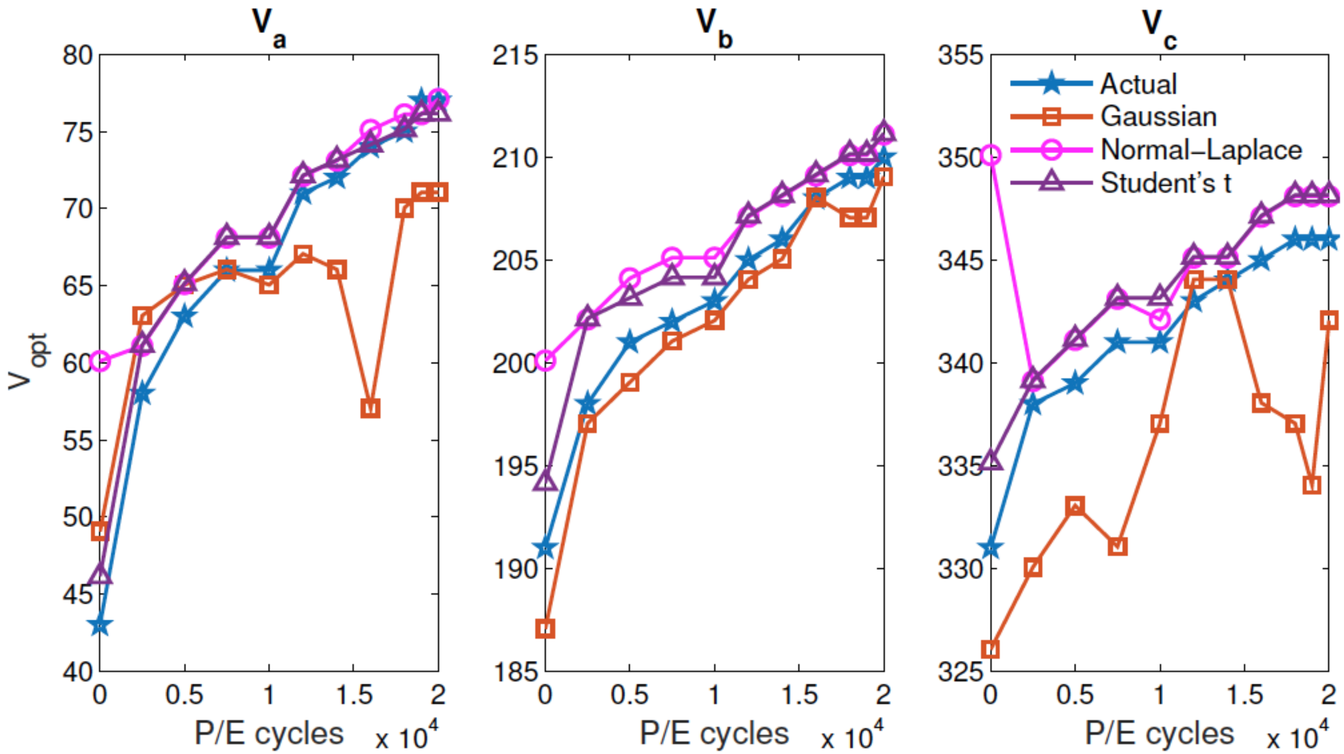


Fig. 16: Actual and modeled *optimal* read reference voltages (V_{opt}) using the three evaluated threshold voltage distribution models at different P/E cycle counts.

Effect on RBER of Read Ref V Prediction

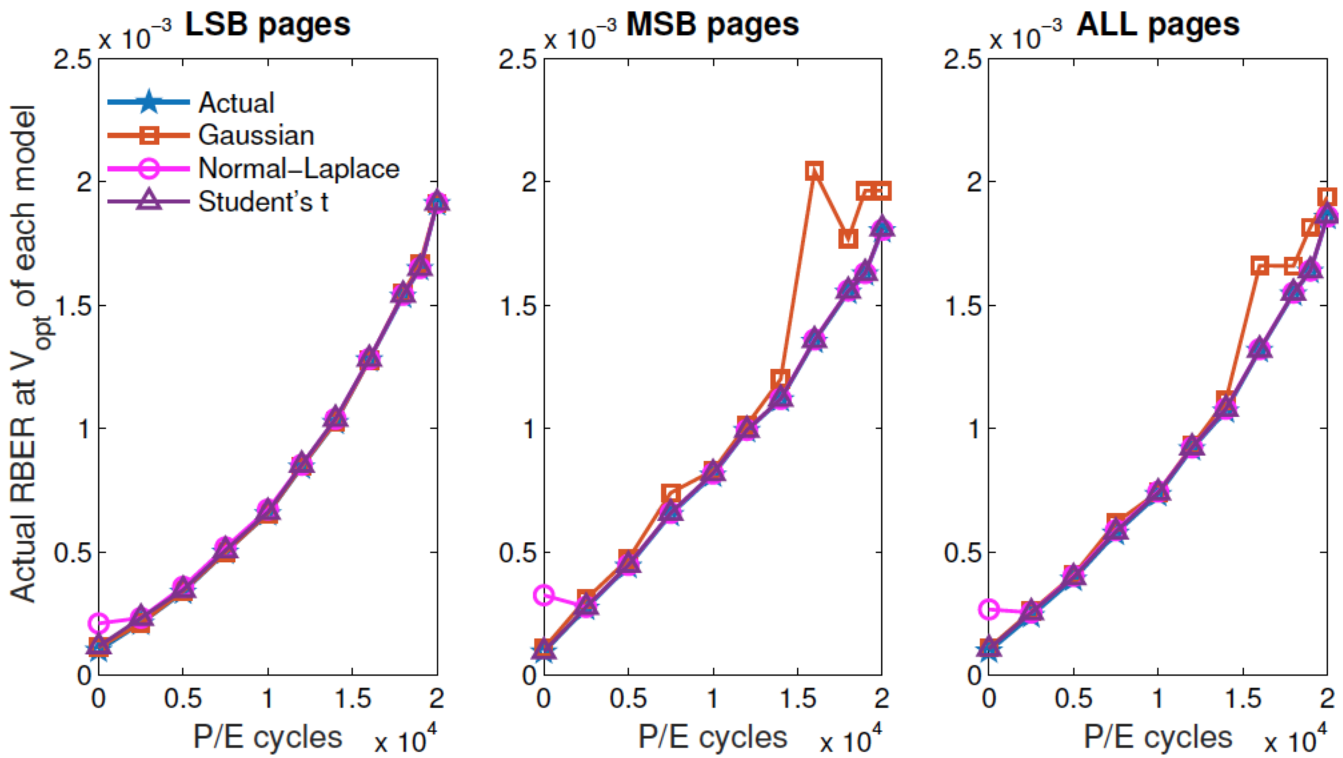


Fig. 17: RBER achieved by actual and modeled *optimal* read reference voltages (V_{opt}) using the three evaluated threshold voltage distribution models at different P/E cycle counts.

More Accurate and Online Channel Modeling

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu, **"Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory"** *to appear in IEEE Journal on Selected Areas in Communications (**JSAC**),* 2016.

Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory

Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, Onur Mutlu

Agenda

- Background, Motivation and Approach
- Experimental Characterization Methodology
- Error Analysis and Management
  - Main Characterization Results
  - Retention-Aware Error Management
  - Threshold Voltage and Program Interference Analysis
  - Read Reference Voltage Prediction
  - Neighbor-Assisted Error Correction
  - Read Disturb Error Handling
  - Retention Error Handling
  - 3D NAND Flash Memory Reliability
- Summary

Goal

- Develop a better error correction mechanism for cases where ECC fails to correct a page

Observations So Far

- Immediate neighbor cell has the most effect on the victim cell when programmed

- A single set of read reference voltages is used to determine the value of the (victim) cell

- The set of read reference voltages is determined based on the ***overall threshold voltage distribution of all cells*** in flash memory

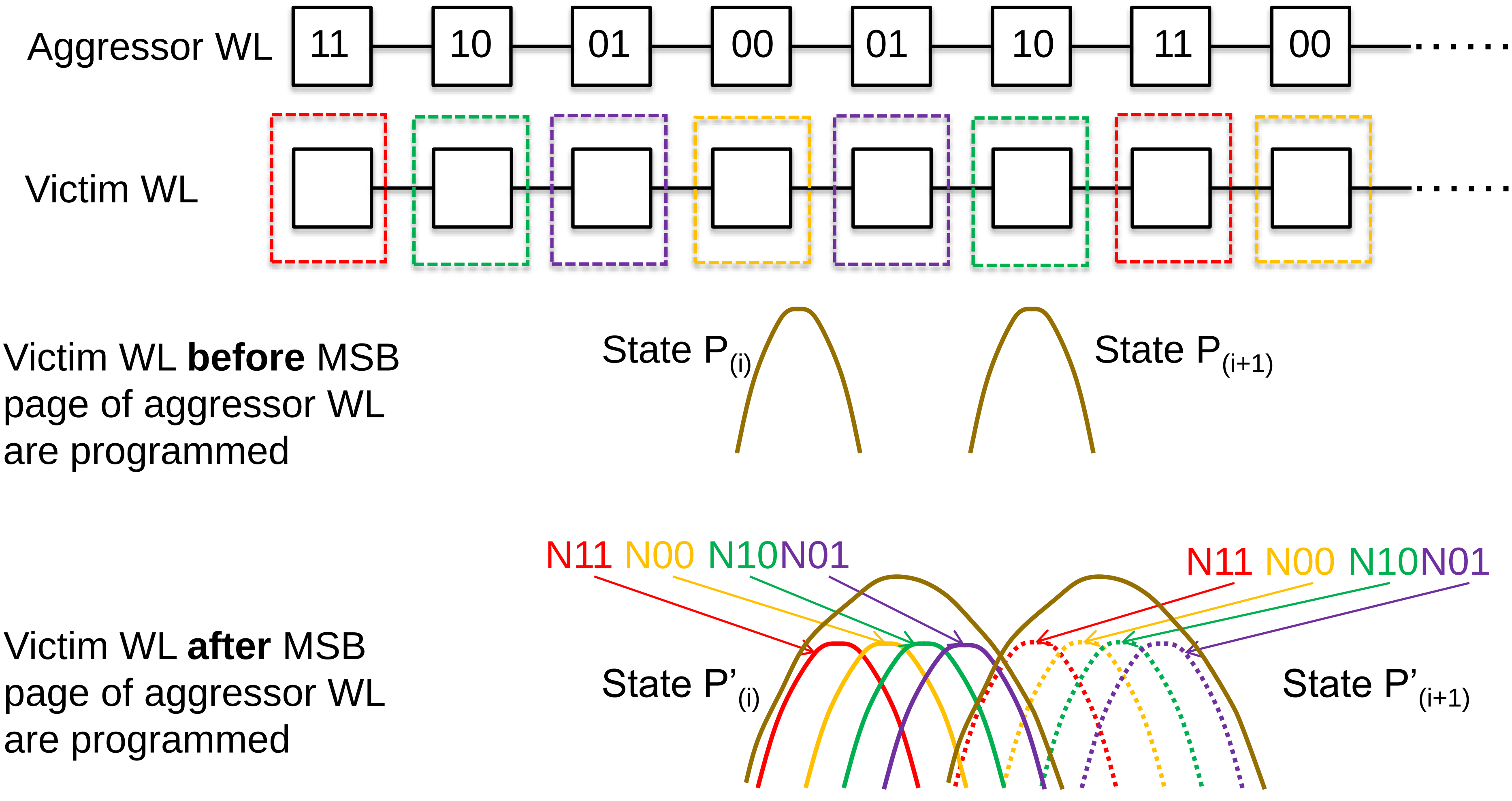
New Observations [Cai+ SIGMETRICS'14]

- Vth distributions of **cells with different-valued immediate-neighbor cells** are significantly different
  - Because neighbor value affects the amount of Vth shift

- Corollary: If we know the value of the immediate-neighbor, we can find a more accurate set of read reference voltages based on the "conditional" threshold voltage distribution

Cai et al., Neighbor-Cell Assisted Error Correction for MLC NAND Flash Memories, SIGMETRICS 2014.

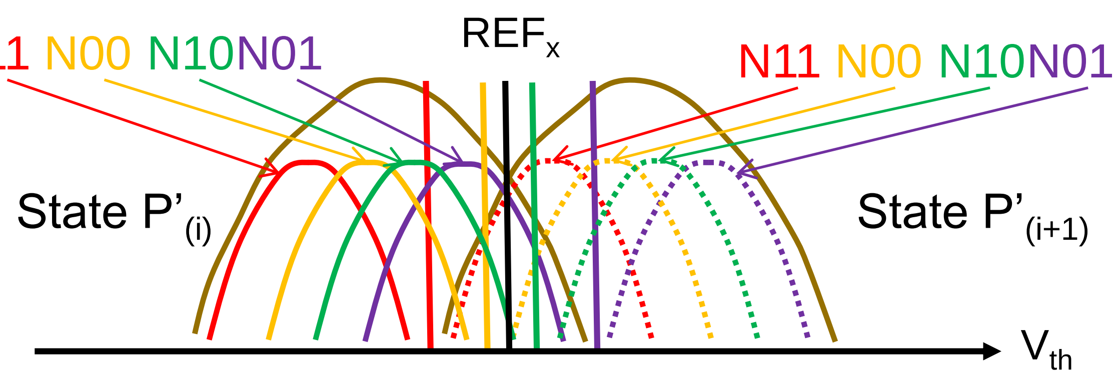
Secrets of Threshold Voltage Distributions



If We Knew the Immediate Neighbor …

- Then, we could choose a different read reference voltage to more accurately read the “victim” cell

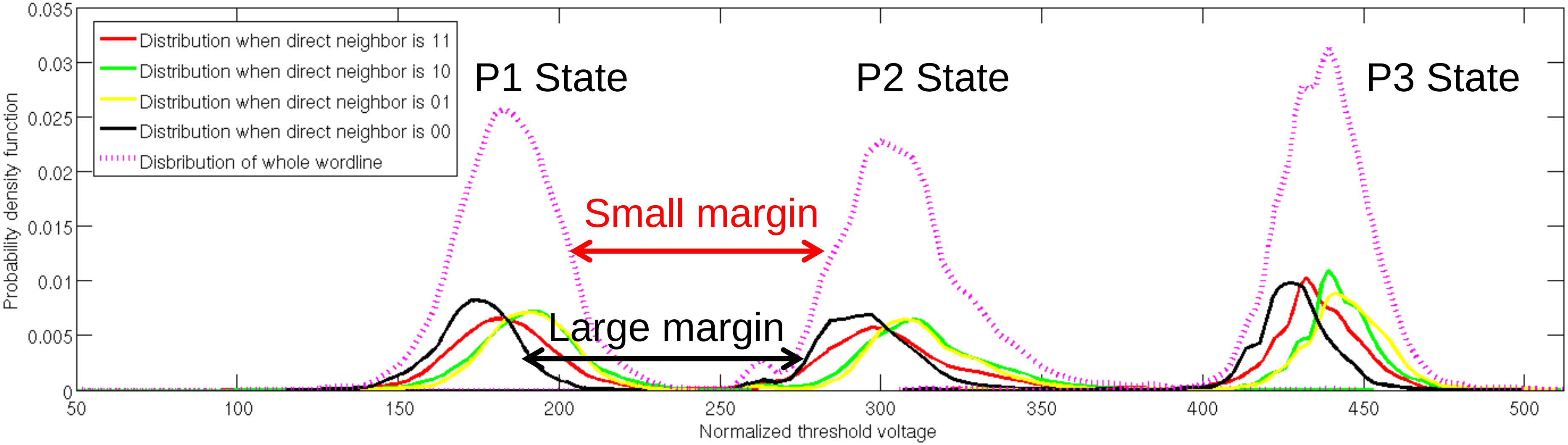
Overall vs Conditional Reading



- Using the optimum read reference voltage based on the overall distribution leads to more errors

- Better to use the optimum read reference voltage based on the conditional distribution (i.e., value of the neighbor)
  - Conditional distributions of two states are farther apart from each other

Real NAND Flash Chip Measurement Results



	Overall	x_{11} (ER)	x_{10} (P1)	x_{00} (P2)	X_{01} (P3)
Distance	65.4	65.4	64.7	66.4	65.8
Variance	385.9	286.2	256.7	242.8	252.1
SNR	3.4	3.8	3.9	4.2	4.1
BER	3×10^{-4}	7×10^{-5}	5×10^{-5}	2×10^{-5}	3×10^{-5}

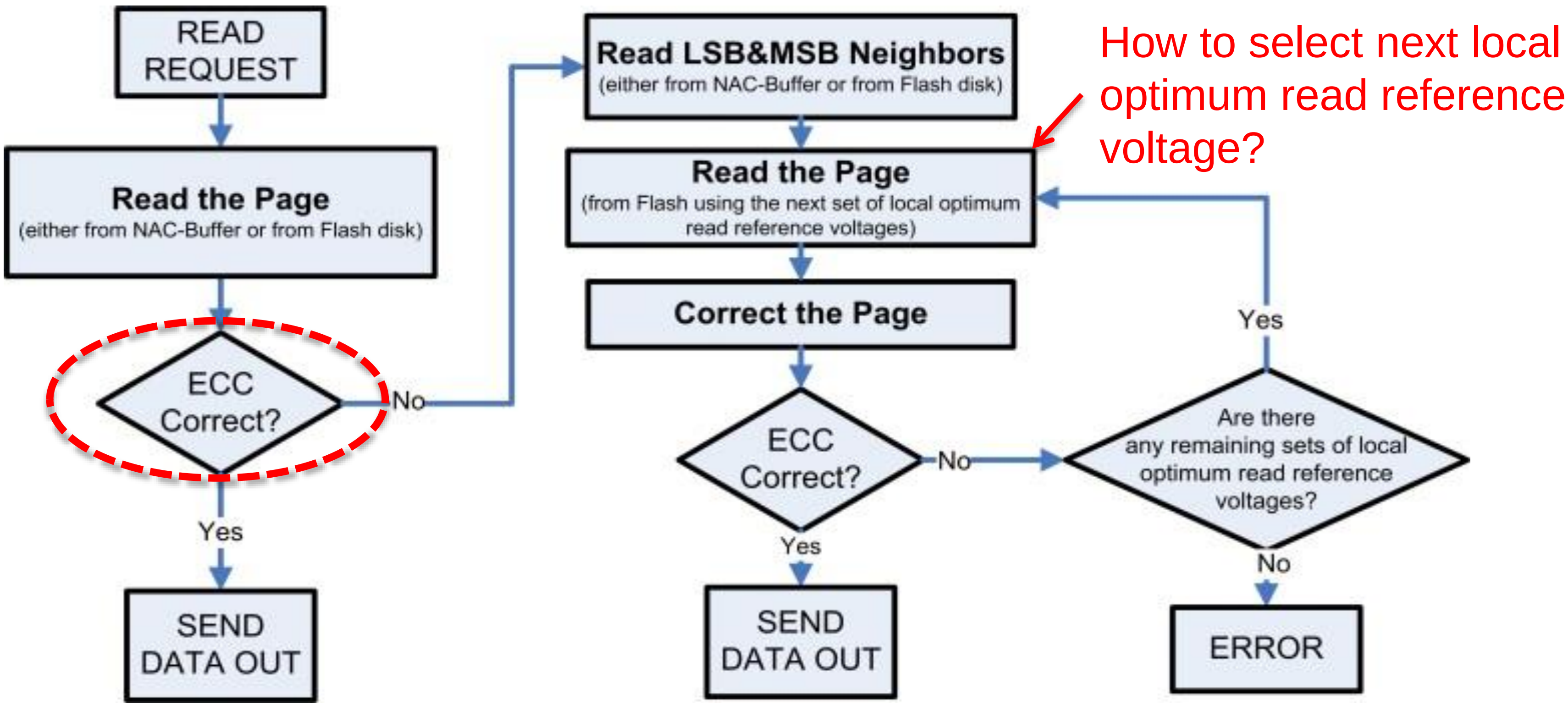
Raw BER of conditional reading is much smaller than overall reading

Idea: Neighbor Assisted Correction (NAC)

- Read a page with the read reference voltages based on overall Vth distribution (same as today) and buffer it

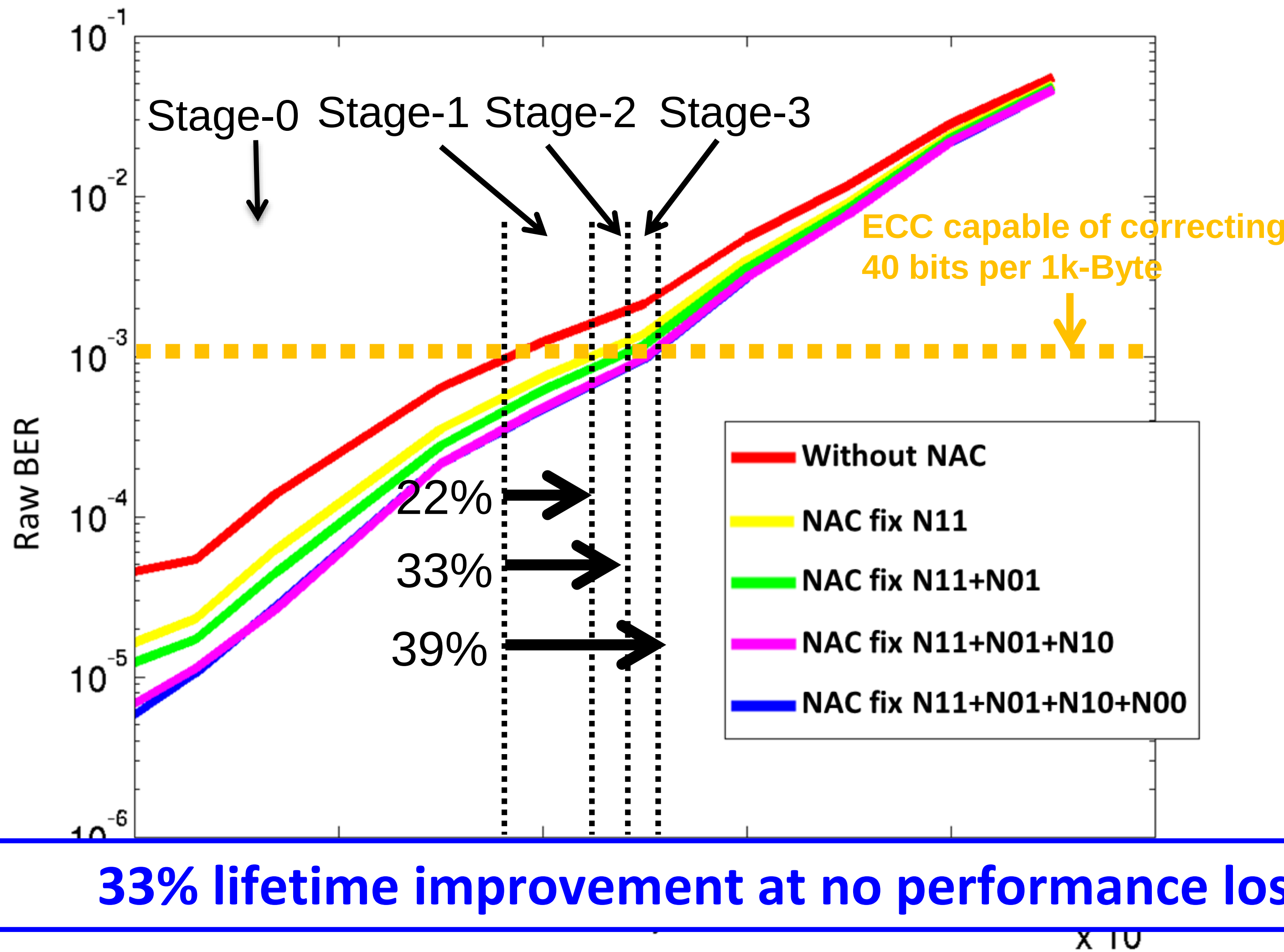
- If ECC fails:
  - Read the immediate-neighbor page
  - Re-read the page using the read reference voltages corresponding to the voltage distribution assuming a particular immediate-neighbor value
  - Replace the buffered values of the cells with that particular immediate-neighbor cell value
  - Apply ECC again

Neighbor Assisted Correction Flow



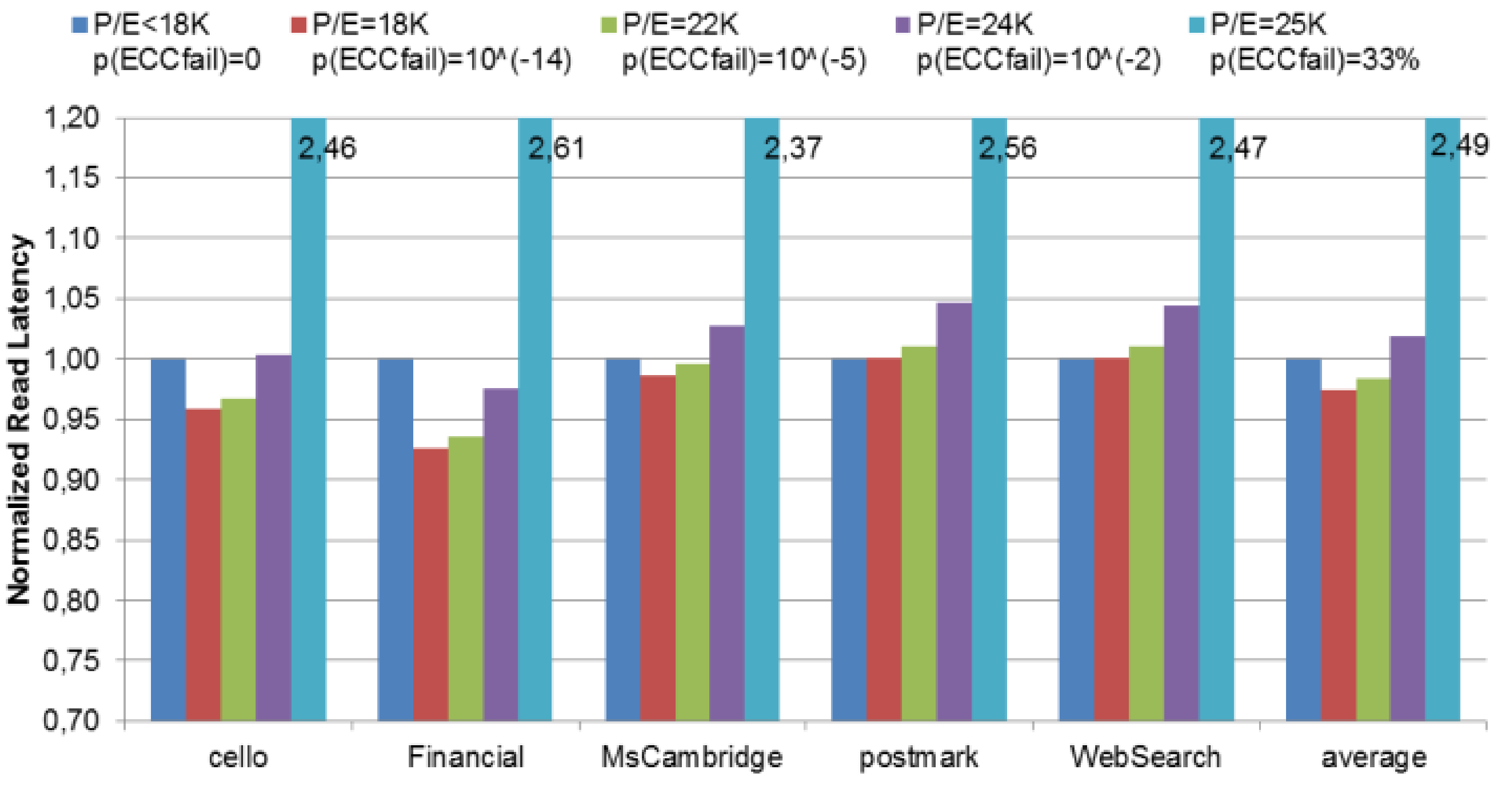
- Trigger neighbor-assisted reading only when ECC fails
- Read neighbor values and use corresponding read reference voltages in a prioritized order until ECC passes

Lifetime Extension with NAC



33% lifetime improvement at no performance loss

Performance Analysis of NAC



No performance loss within nominal lifetime and with reasonable (1%) ECC fail rates

More on Neighbor-Assisted Correction

- Yu Cai, Gulay Yalcin, Onur Mutlu, Eric Haratsch, Osman Unsal, Adrian Cristal, and Ken Mai,
"Neighbor-Cell Assisted Error Correction for MLC NAND Flash Memories"
Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (***SIGMETRICS***), Austin, TX, June 2014. Slides (ppt) (pdf)

Neighbor-Cell Assisted Error Correction for MLC NAND Flash Memories

Yu Cai[1], Gulay Yalcin[2], Onur Mutlu[1], Erich F. Haratsch[4],
Osman Unsal[2], Adrian Cristal[2,3], and Ken Mai[1]
[1]Electrical and Computer Engineering Department, Carnegie Mellon University
[2]Barcelona Supercomputing Center, Spain [3]IIIA – CSIC – Spain National Research Council [4]LSI Corporation
yucaicai@gmail.com, {omutlu, kenmai}@ece.cmu.edu, {gulay.yalcin, adrian.cristal, osman.unsal}@bsc.es

Agenda

- Background, Motivation and Approach
- Experimental Characterization Methodology
- Error Analysis and Management
  - Main Characterization Results
  - Retention-Aware Error Management
  - Threshold Voltage and Program Interference Analysis
  - Read Reference Voltage Prediction
  - Neighbor-Assisted Error Correction
  - Read Disturb Error Handling
  - Retention Error Handling
  - 3D NAND Flash Memory Reliability
- Summary

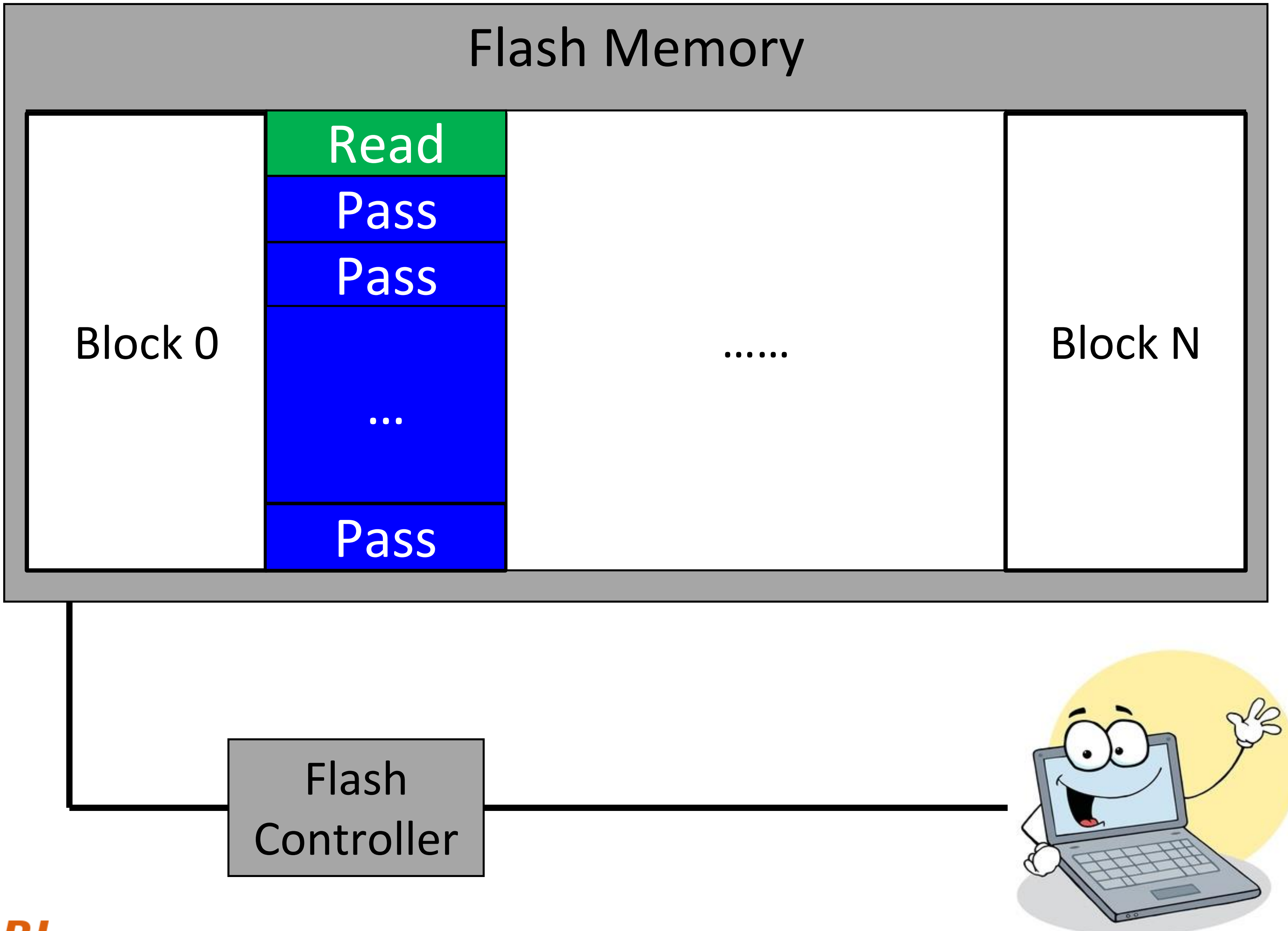
Read Disturb Errors in Flash Memory

One Issue: Read Disturb in Flash Memory

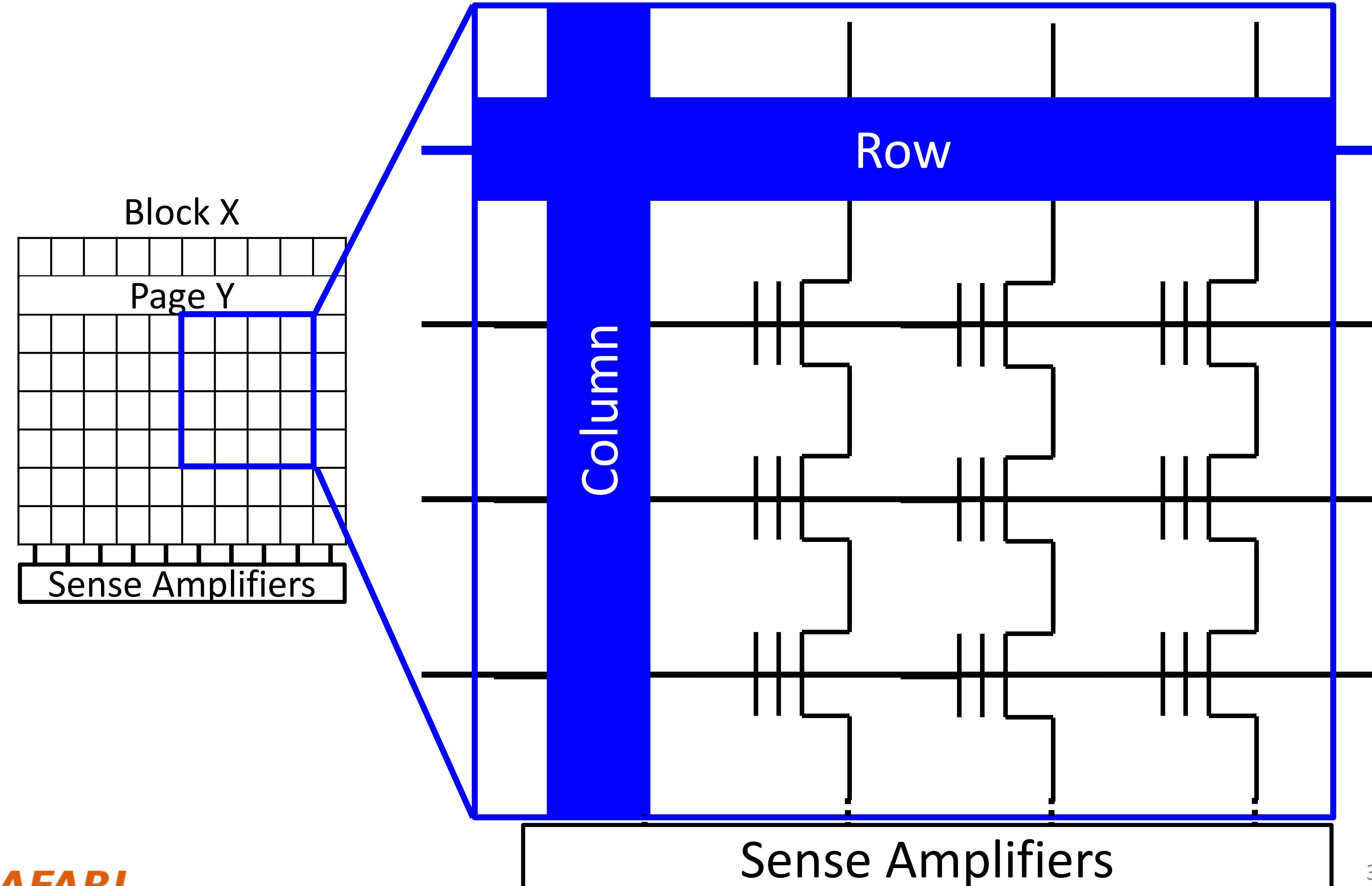
- All scaled memories are prone to read disturb errors

- DRAM
- SRAM
- Hard Disks: Adjacent Track Interference
- NAND Flash

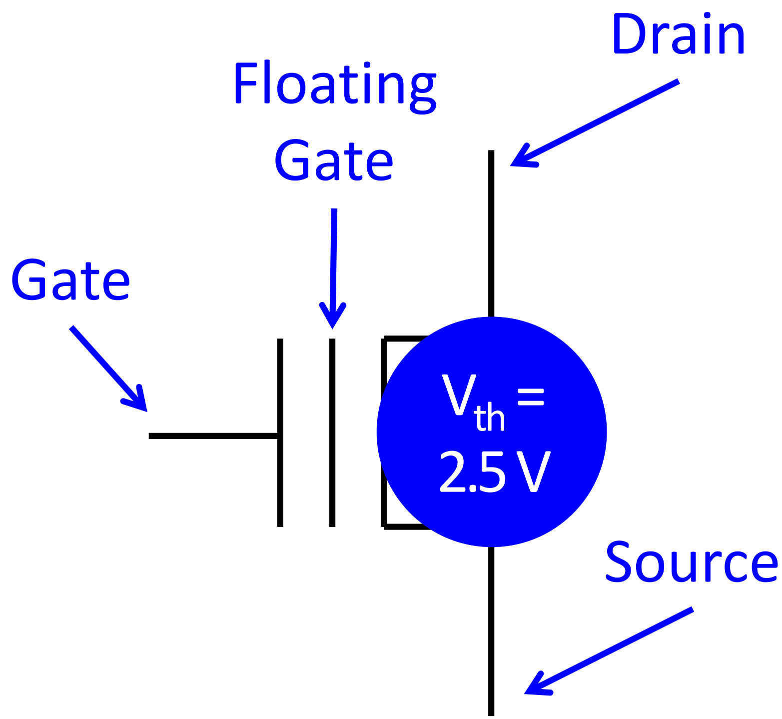
NAND Flash Memory Background



Flash Cell Array

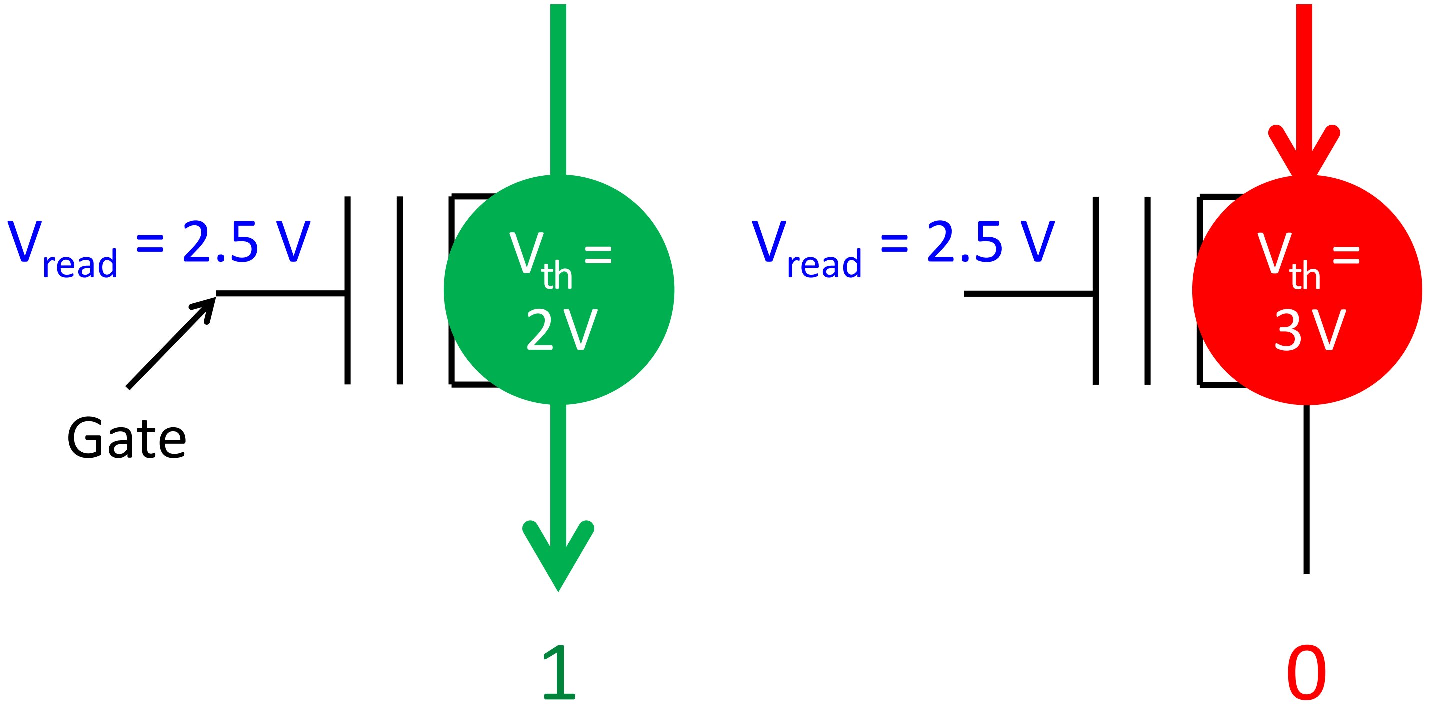


Flash Cell

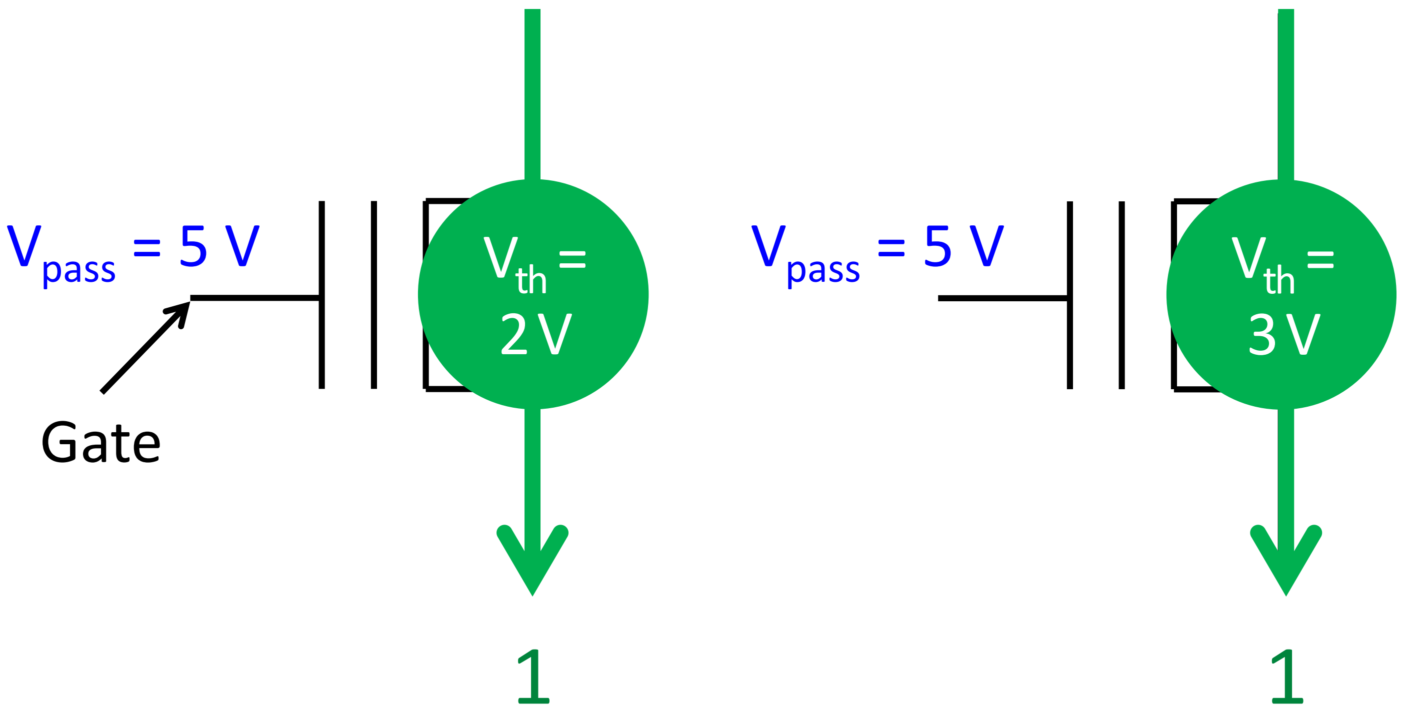


Floating Gate Transistor
(Flash Cell)

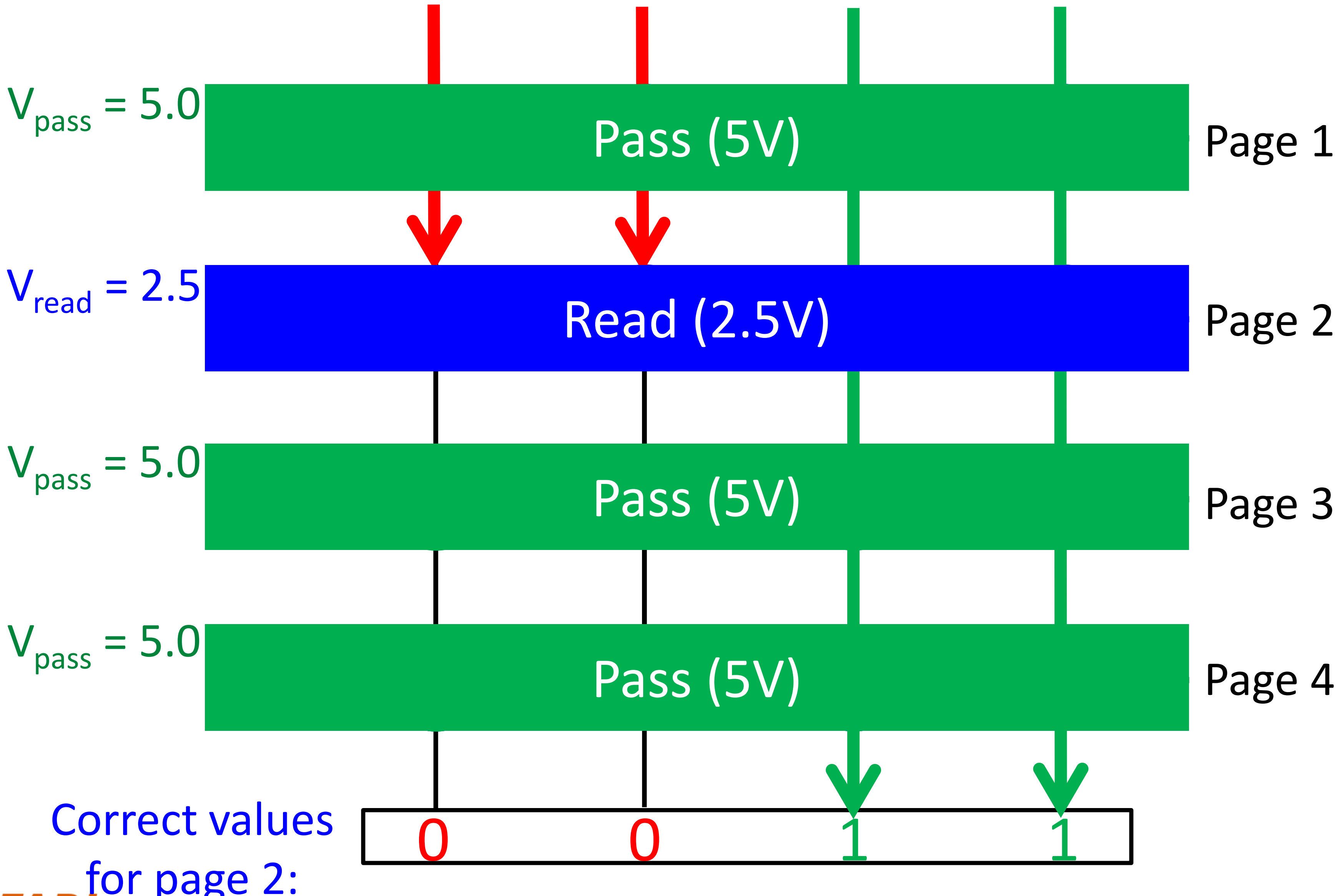
Flash Read



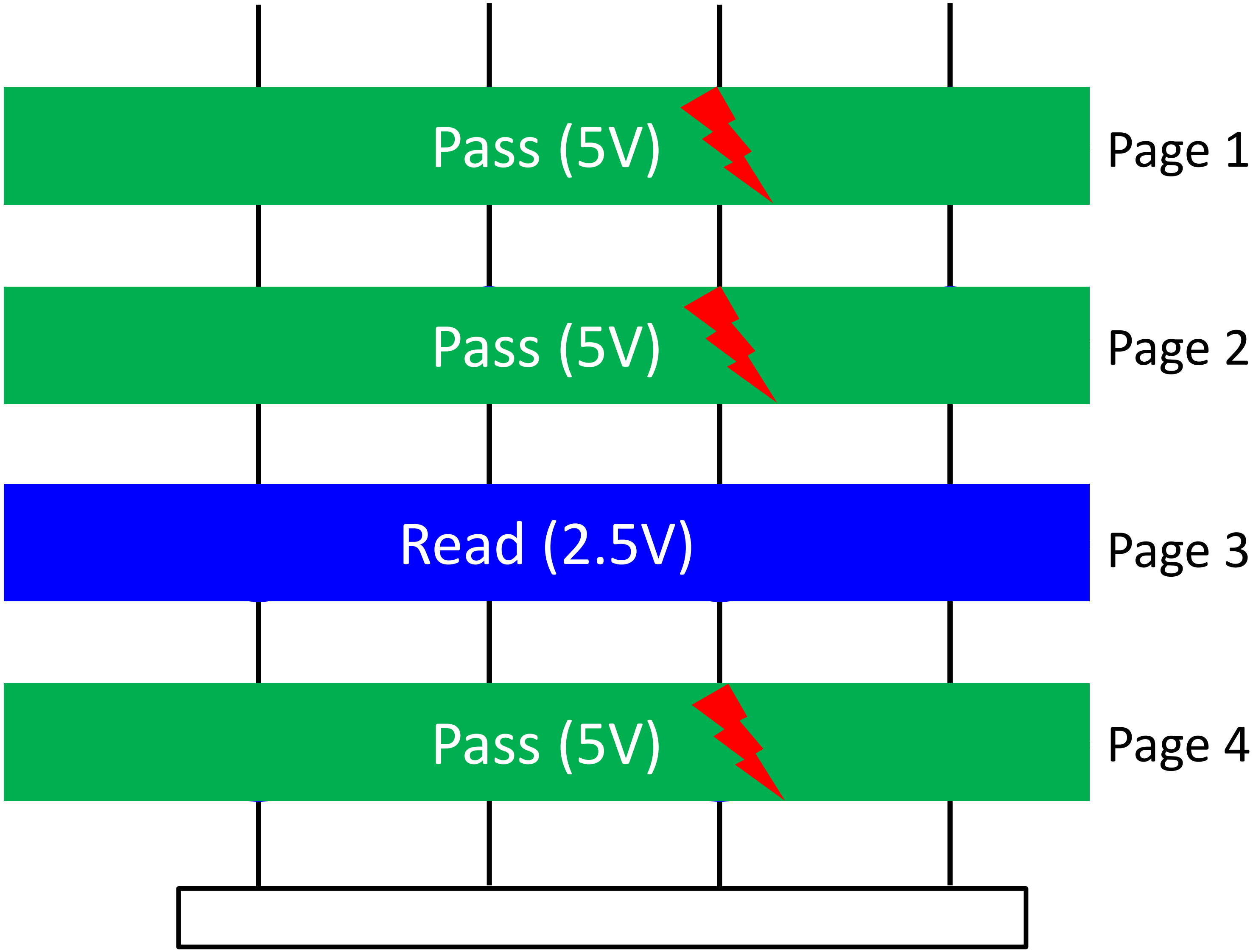
Flash Pass-Through



Read from Flash Cell Array

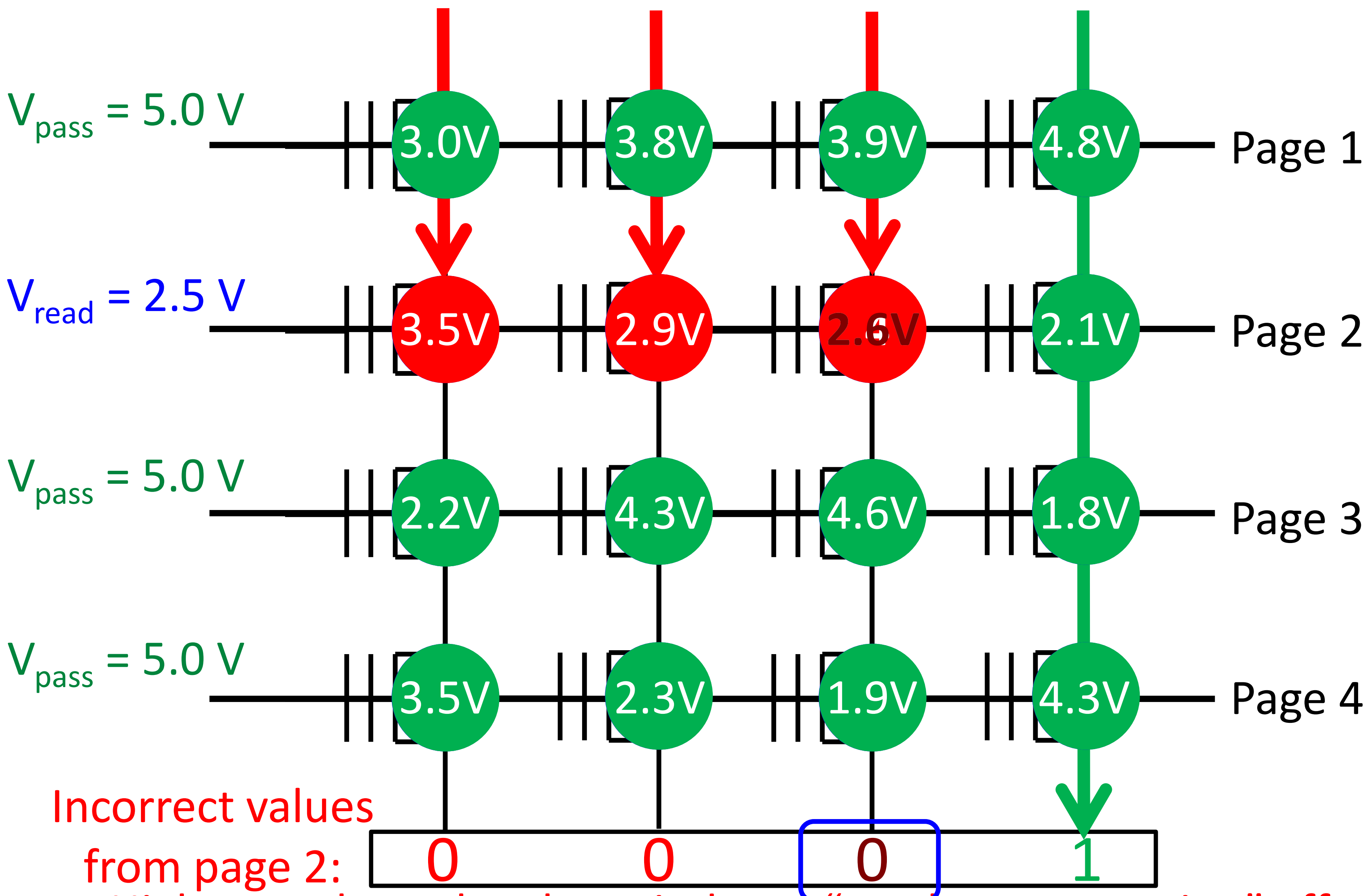


Read Disturb Problem: "Weak Programming" Effect



Repeatedly read page 3 (or any page other than page 2)

Read Disturb Problem: “Weak Programming” Effect



High pass-through voltage induces “weak-programming” effect

Executive Summary [DSN'15]

- ***Read disturb errors*** limit flash memory lifetime today
  - Apply a *high pass-through voltage (V_{pass})* to multiple pages on a read
  - Repeated application of V_{pass} can alter stored values in unread pages
- We **characterize read disturb** on real NAND flash chips
  - Slightly lowering V_{pass} greatly reduces read disturb errors
  - Some flash cells are more prone to read disturb
- **Technique 1:** Mitigate read disturb errors online
  - ***V_{pass} Tuning*** dynamically finds and applies a lowered V_{pass} per block
  - Flash memory lifetime improves by 21%
- **Technique 2:** Recover after failure to prevent data loss
  - ***Read Disturb Oriented Error Recovery*** (RDR) selectively corrects cells more susceptible to read disturb errors
  - Reduces raw bit error rate (RBER) by up to 36%

Key Observation 1: Slightly lowering V_{pass} greatly reduces read disturb errors

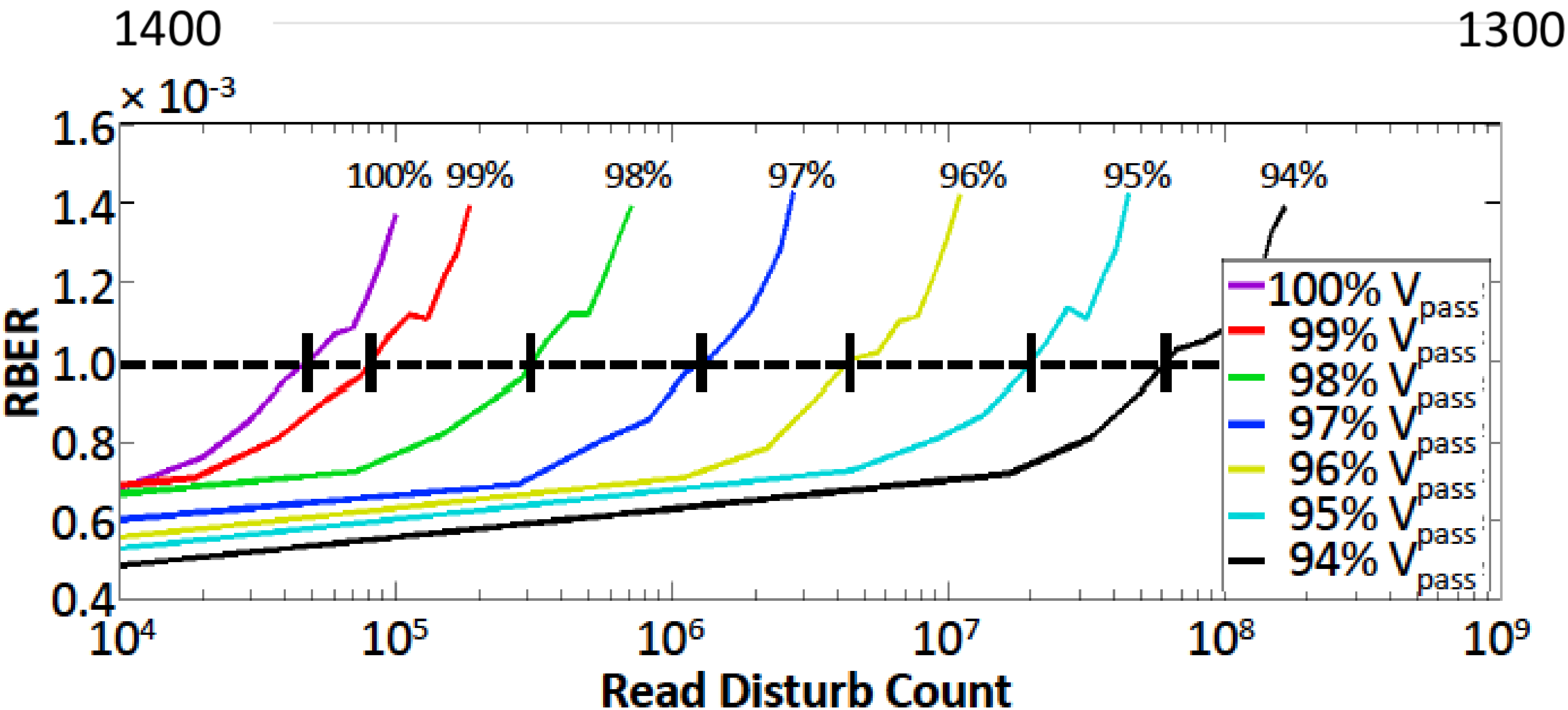


Fig. 11. Raw bit error rate vs. read disturb count for different V_{pass} values, for flash memory under 8K P/E cycles of wear.

Percentage of Vpass Reduction

Outline

- Background (Problem and Goal)
- Key Experimental Observations
- Mitigation: V_{pass} Tuning
- Recovery: Read Disturb Oriented Error Recovery
- Conclusion

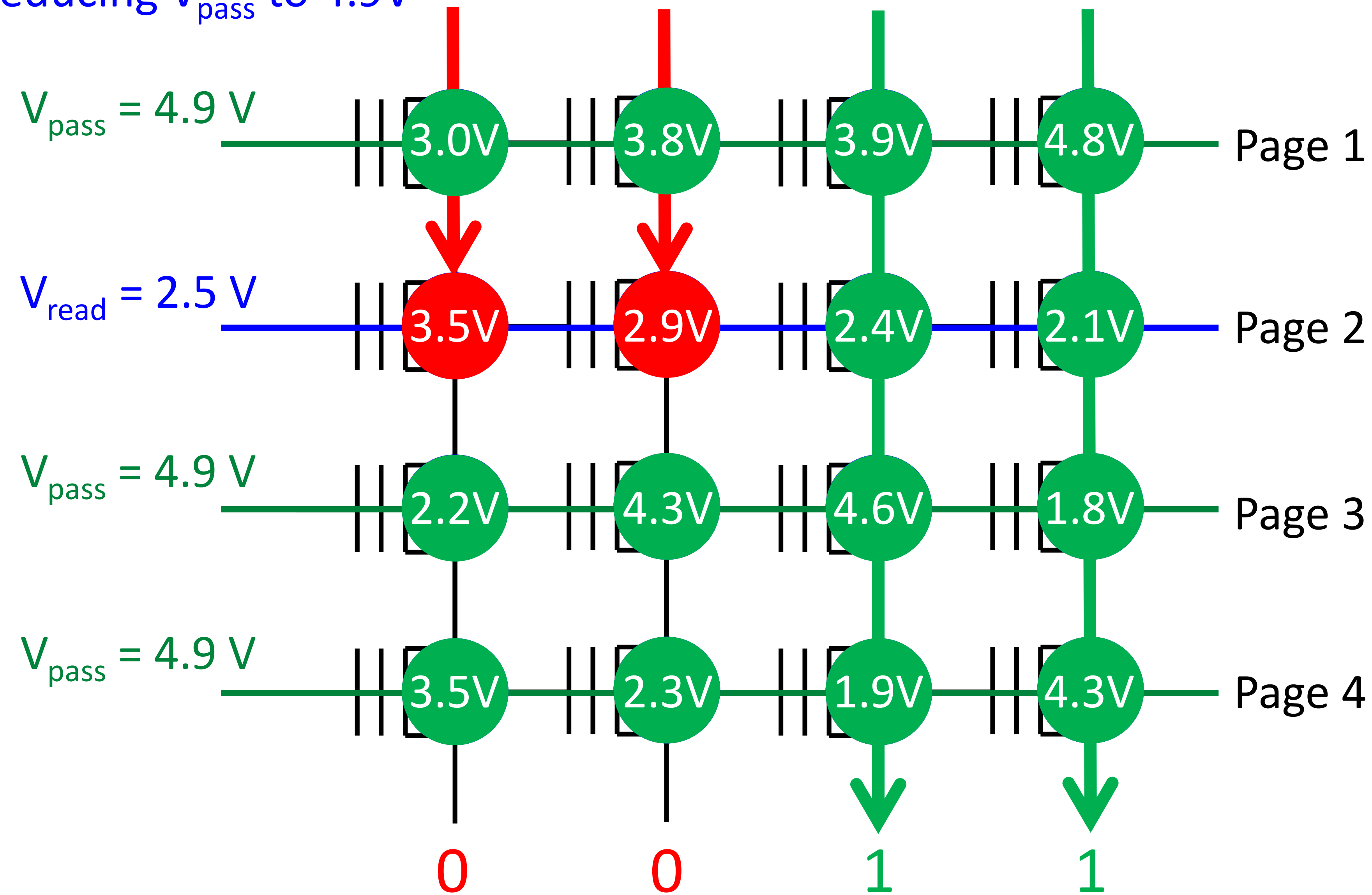
Read Disturb Mitigation: V_{pass} Tuning

- Key Idea: Dynamically find and apply a lowered V_{pass}

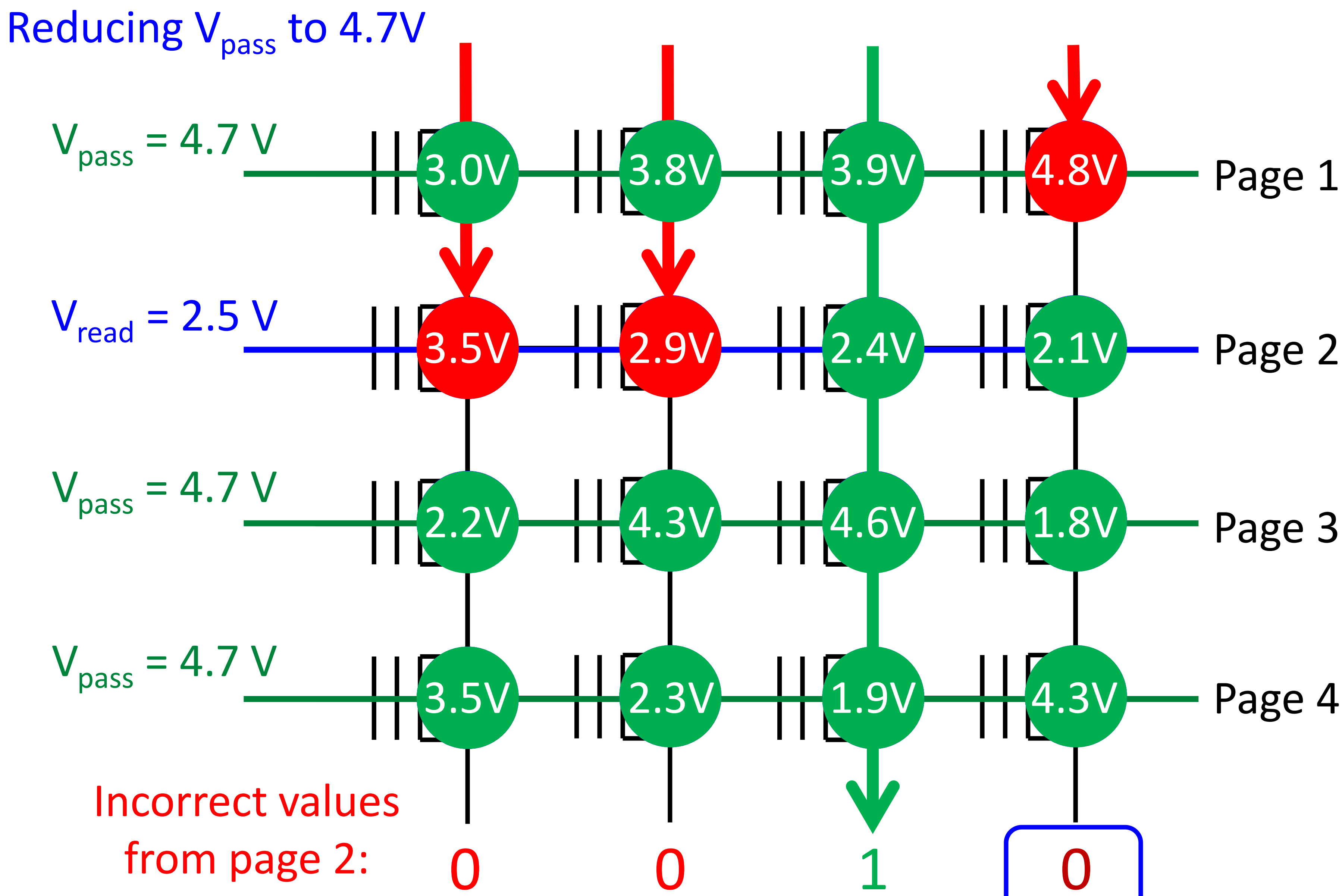
- Trade-off for lowering V_{pass}
  + Allows more read disturbs
  – Induces more read errors

Read Errors Induced by V_{pass} Reduction

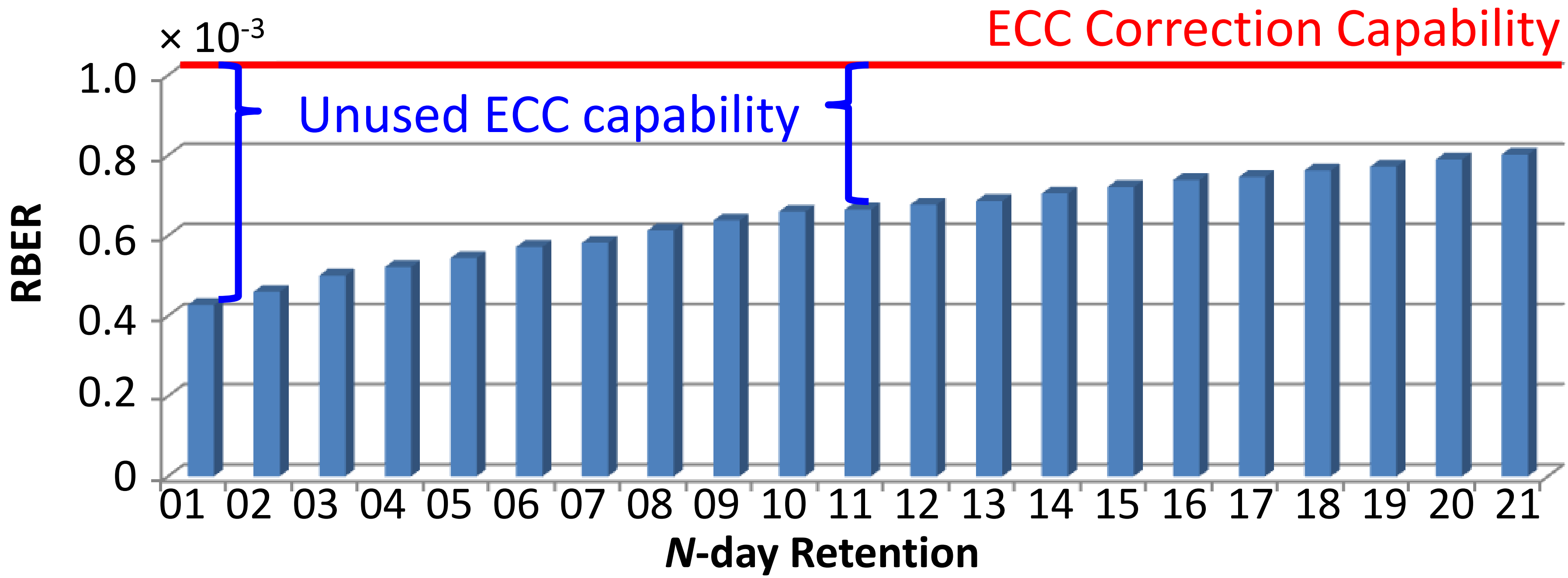
Reducing V_{pass} to 4.9V



Read Errors Induced by V_{pass} Reduction



Utilizing the Unused ECC Capability



1. ECC provisioned for high retention "age"
2. Unused ECC capability can be used to fix read errors
3. Unused ECC capability decreases over retention age

Dynamically adjust V_{pass} so that read errors fully utilize the unused ECC capability

V_{pass} Reduction Trade-Off Summary

- Today: Conservatively set V_{pass} to a high voltage
  - – Accumulates more read disturb errors at the end of each refresh interval
  - \+ No read errors

- Idea: Dynamically adjust V_{pass} to unused ECC capability
  - \+ Minimize read disturb errors
  - o Control read errors to be tolerable by ECC
  - o If read errors exceed ECC capability, read again with a higher V_{pass} to correct read errors

V_{pass} Tuning Steps

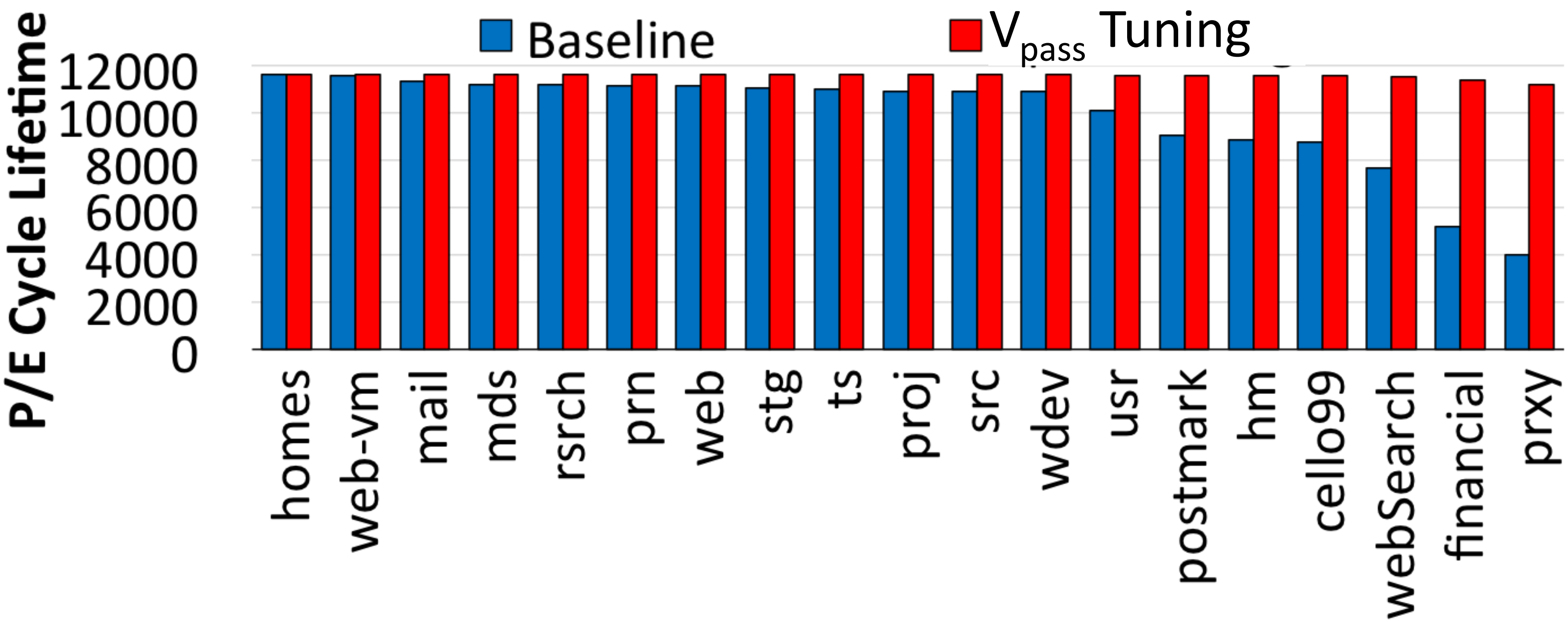
- Perform once for each block every day:
  1. Estimate *unused ECC capability (using retention age)*
  2. Aggressively reduce V_{pass} until *read errors exceeds ECC capability*
  3. Gradually increase V_{pass} until read error becomes just less than ECC capability

Evaluation of V_{pass} Tuning

- 19 real workload I/O traces
- Assume 7-day refresh period
- Similar methodology as before to determine acceptable V_{pass} reduction

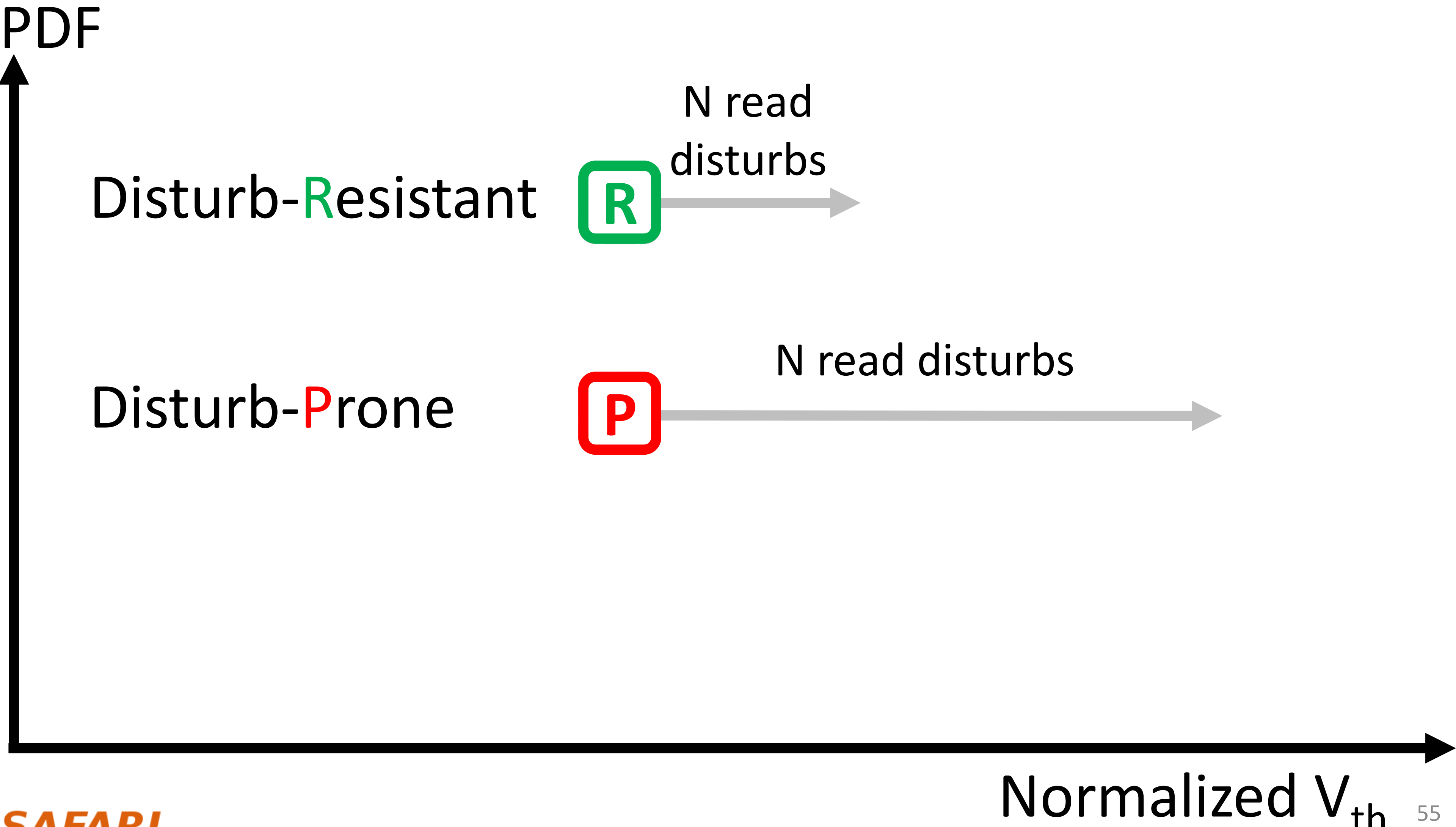
- Overhead for a 512 GB flash drive:
  - 128 KB storage overhead for per-block V_{pass} setting and worst-case page
  - 24.34 sec/day average V_{pass} Tuning overhead

V_{pass} Tuning Lifetime Improvements



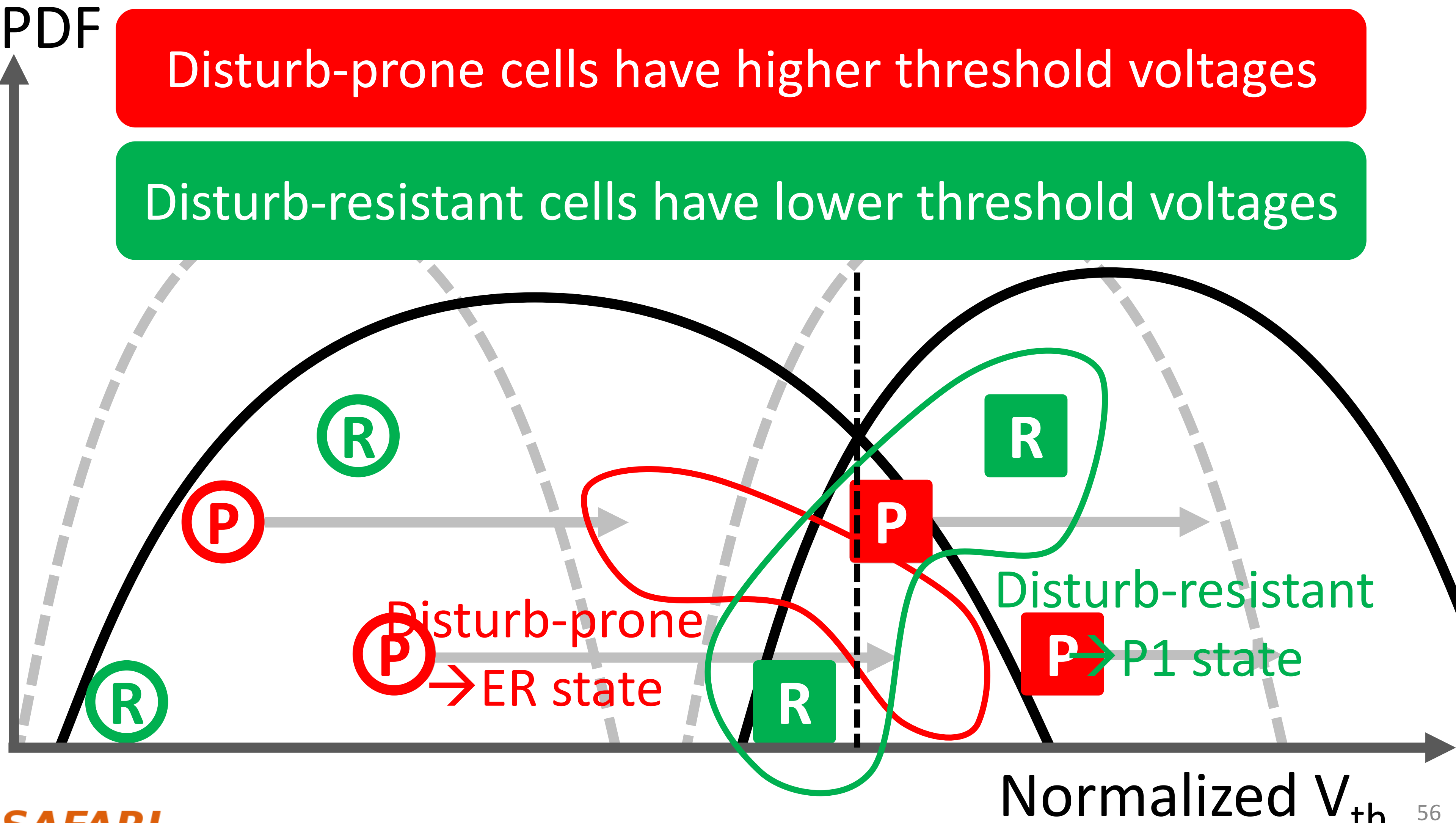
Average lifetime improvement: 21.0%

Read Disturb Prone vs. Resistant Cells



Observation 2: Some Flash Cells Are More Prone to Read Disturb

After 250K read disturbs:

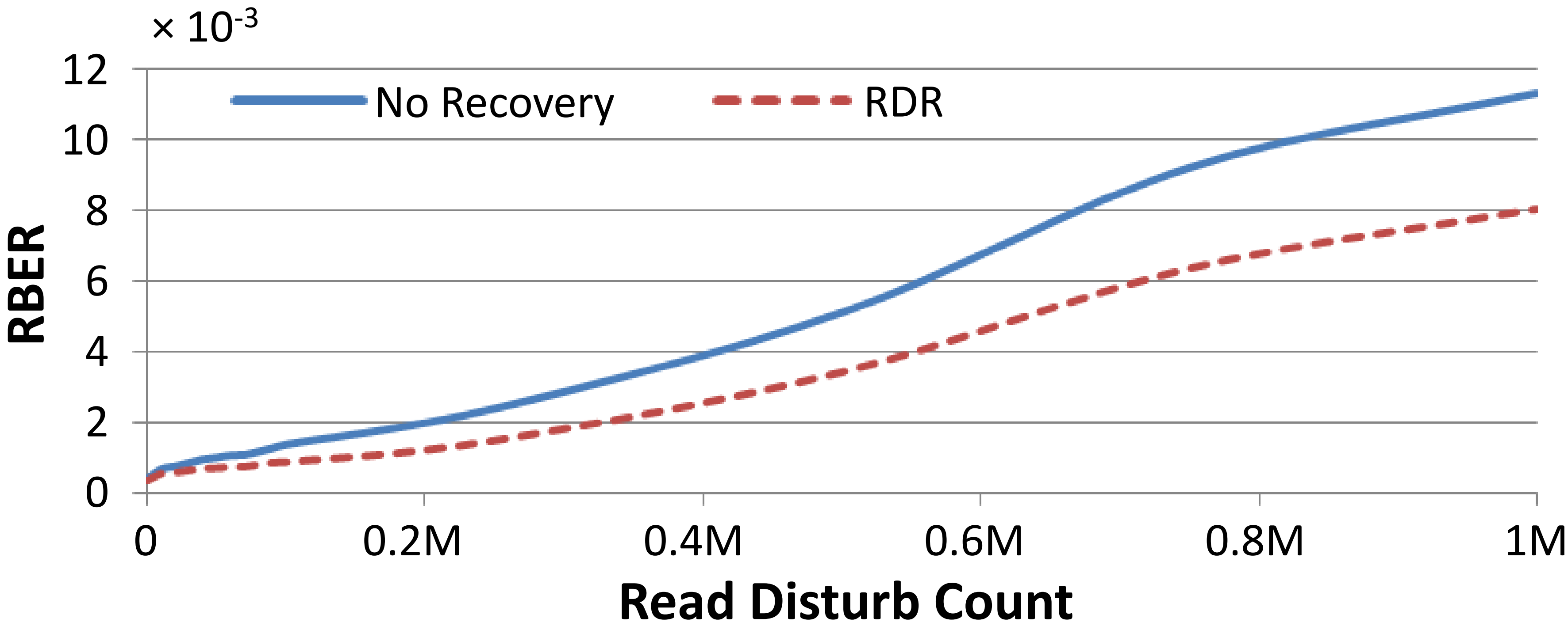


Read Disturb Oriented Error Recovery (RDR)

- Triggered by an uncorrectable flash error
  - Back up all valid data in the faulty block
  - Disturb the faulty page 100K times (more)
  - Compare V_{th}'s before and after read disturb
  - Select cells susceptible to flash errors ($V_{ref}-\sigma<V_{th}<V_{ref}-\sigma$)
  - Predict among these susceptible cells
    - Cells with more V_{th} shifts are disturb-prone → Lower V_{th} state
    - Cells with less V_{th} shifts are disturb-resistant → Higher V_{th} state

Reduces total error count by up to 36% @ 1M read disturbs
ECC can be used to correct the remaining errors

RDR Evaluation



Reduces total error counts by up to 36% @ 1M read disturbs
ECC can be used to correct the remaining errors

More on Flash Read Disturb Errors [DSN'15]

- Yu Cai, Yixin Luo, Saugata Ghose, Erich F. Haratsch, Ken Mai, and Onur Mutlu,
"Read Disturb Errors in MLC NAND Flash Memory: Characterization and Mitigation"
*Proceedings of the 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (**DSN**)*, Rio de Janeiro, Brazil, June 2015.

Read Disturb Errors in MLC NAND Flash Memory: Characterization, Mitigation, and Recovery

Yu Cai, Yixin Luo, Saugata Ghose, Erich F. Haratsch*, Ken Mai, Onur Mutlu
Carnegie Mellon University, *Seagate Technology
yucaicai@gmail.com, {yixinluo, ghose, kenmai, onur}@cmu.edu

Agenda

- Background, Motivation and Approach
- Experimental Characterization Methodology
- Error Analysis and Management
  - Main Characterization Results
  - Retention-Aware Error Management
  - Threshold Voltage and Program Interference Analysis
  - Read Reference Voltage Prediction
  - Neighbor-Assisted Error Correction
  - Read Disturb Error Handling
  - Retention Error Handling
  - 3D NAND Flash Memory Reliability
- Summary

Data Retention in Flash Memory

Characterize retention loss in real NAND chip

Optimize read performance for old data

Recover old data after failure

TECHSPOT

TRENDING ▾ REVIEWS ▾ FEATURES ▾ DOWNLOADS ▾ PRODUCT FINDER ▾ FORUMS ▾ TEC

An unfortunate tale about Samsung's SSD 840 read performance degradation

An avalanche of reports emerged last September, when owners of the usually speedy Samsung SSD 840 and SSD 840 EVO detected the drives were no longer performing as they used to.

The issue has to do with older blocks of data: reading **old files** nsistently slower than norma **as slow as 30MB/s** whereas **newly-written files** ones used in benchmarks, perform as fast as new – aro **500 MB/s** for the well regarded SSD 840 EVO. The reason no one had noticed (we reviewed the drive back in September 2013) is that data has to be several weeks old to show the problem. Samsung promptly admitted the issue and proposed a fix.

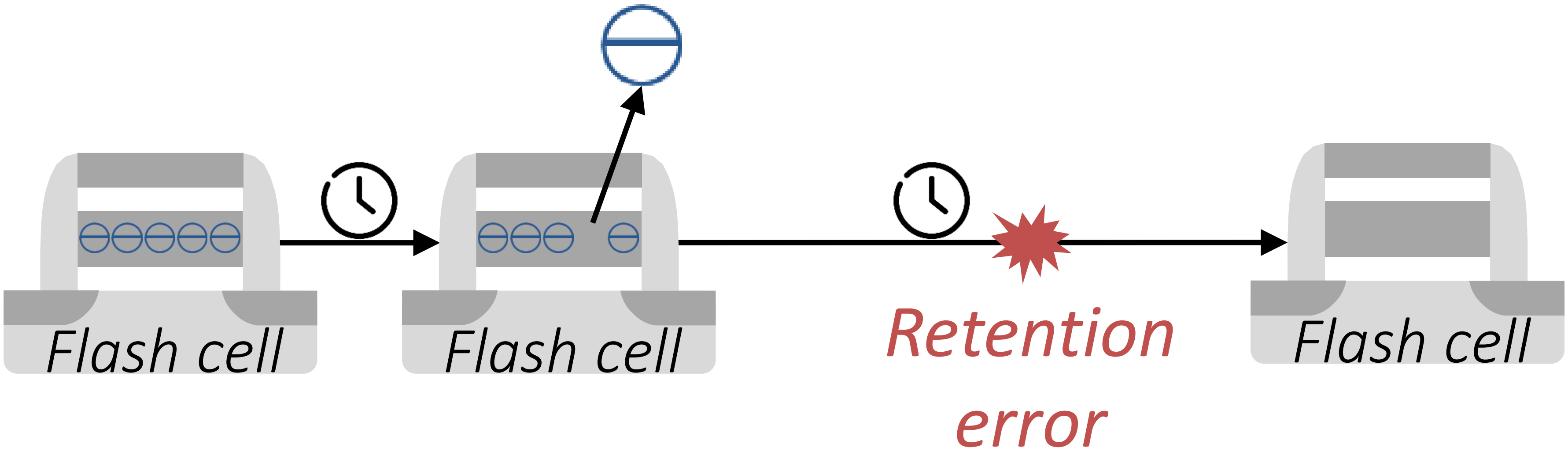
Reference: (May 5, 2015) Per Hansson, "*When SSD Performance Goes Awry*"
http://www.techspot.com/article/997-samsung-ssd-read-performance-degradation/

Why is old data slower?

Retention loss!

Retention loss

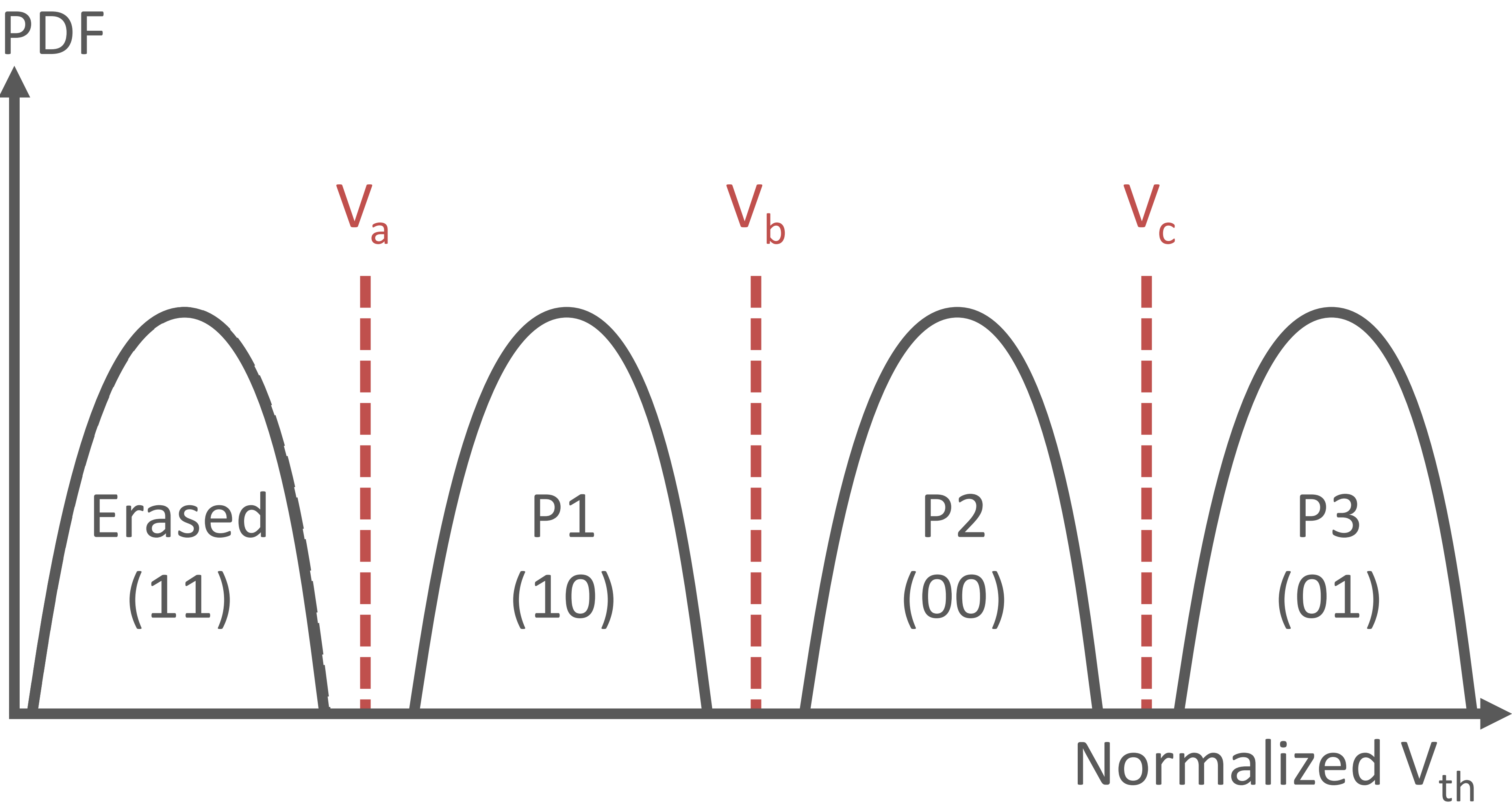
Charge leakage over time



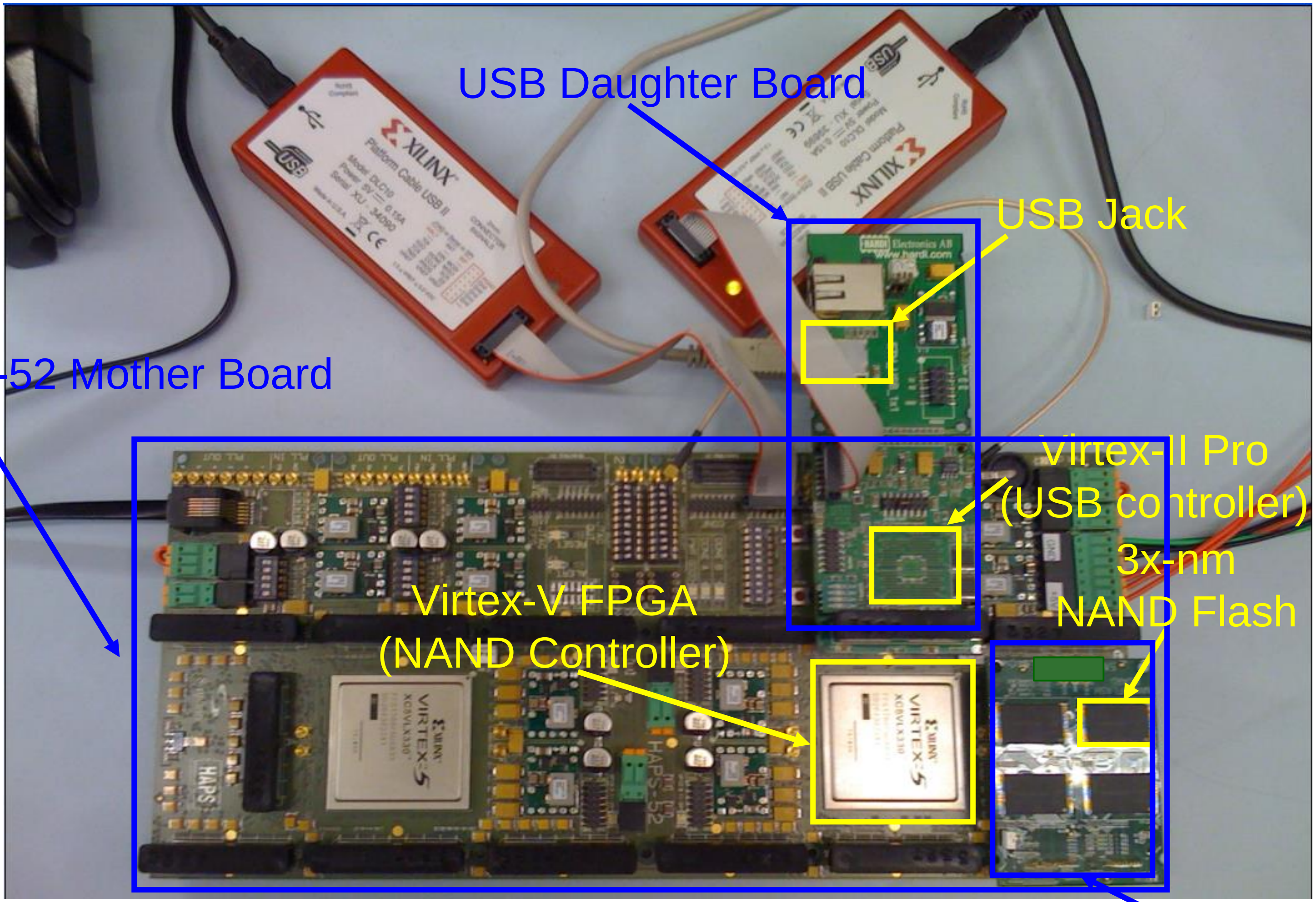
One dominant source of flash memory errors [DATE '12, ICCD '12]

Side effect: Longer read latency

Multi-Level Cell (MLC) threshold voltage distribution



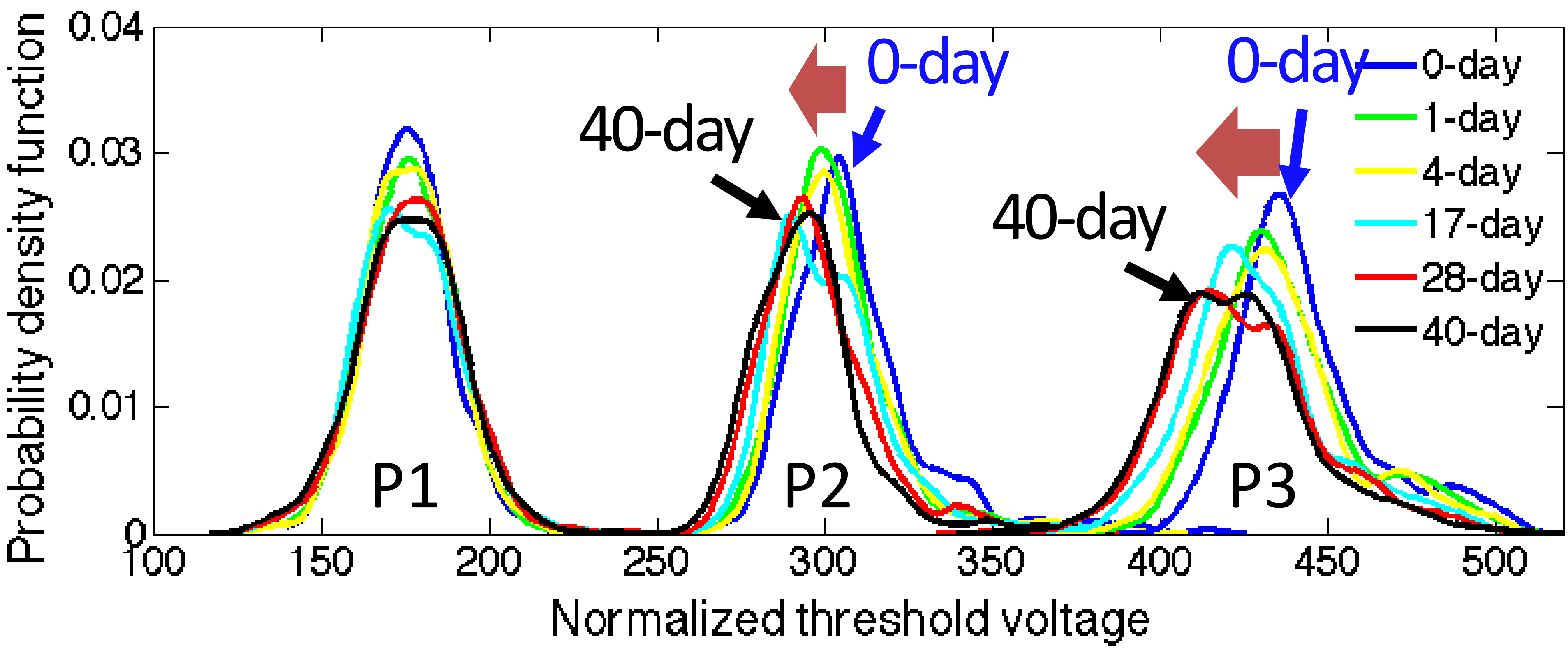
Experimental Testing Platform



[Cai+, FCCM 2011, DATE 2012, ICCD 2012, DATE 2013, ITJ 2013, ICCD 2013, SIGMETRICS 2014, DSN 2015, HPCA 2015]

Cai et al., FPGA-based Solid-State Drive prototyping platform, FCCM 2011.

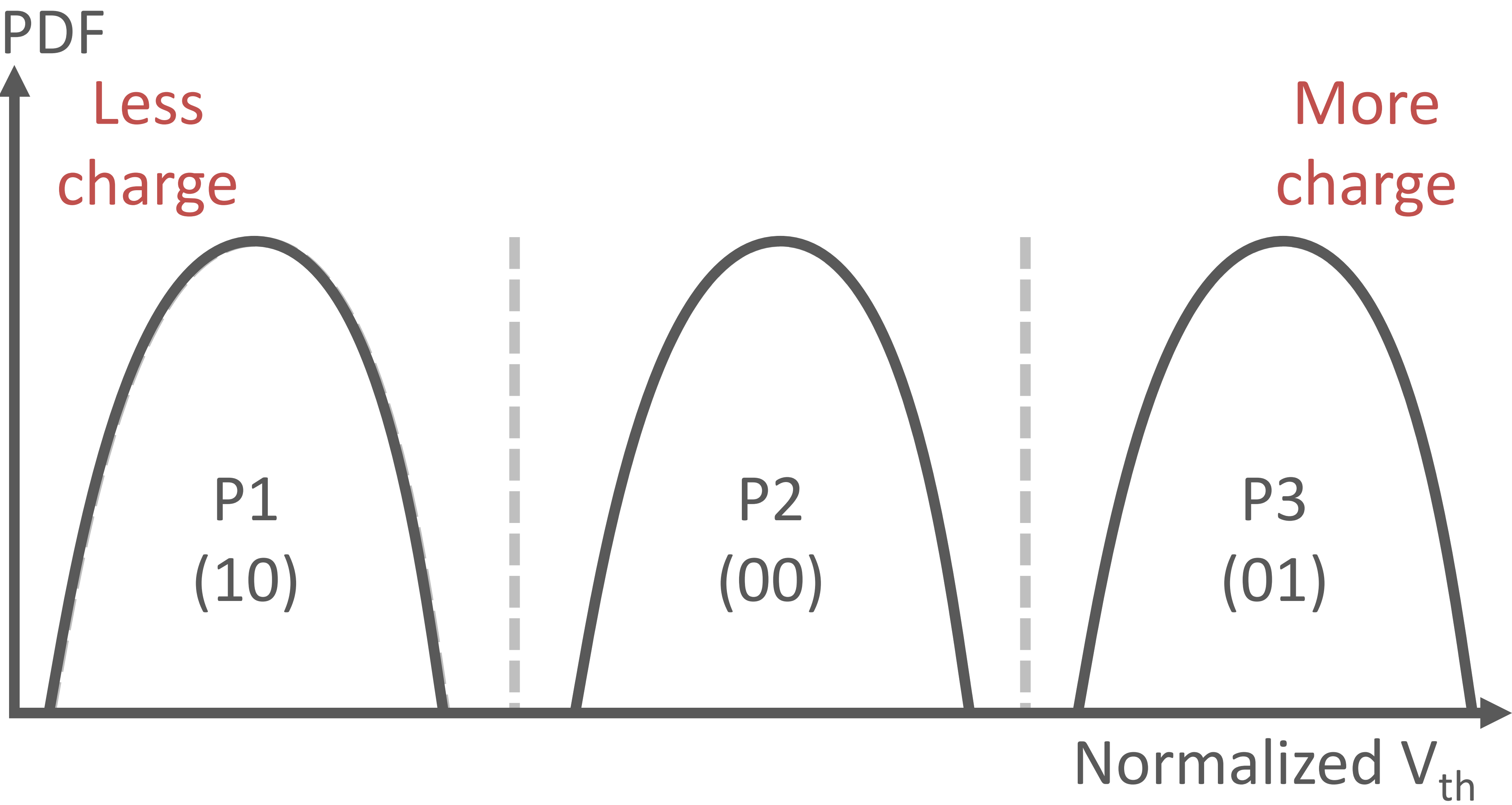
Characterized threshold voltage distribution



Finding: Cell's threshold voltage decreases over time

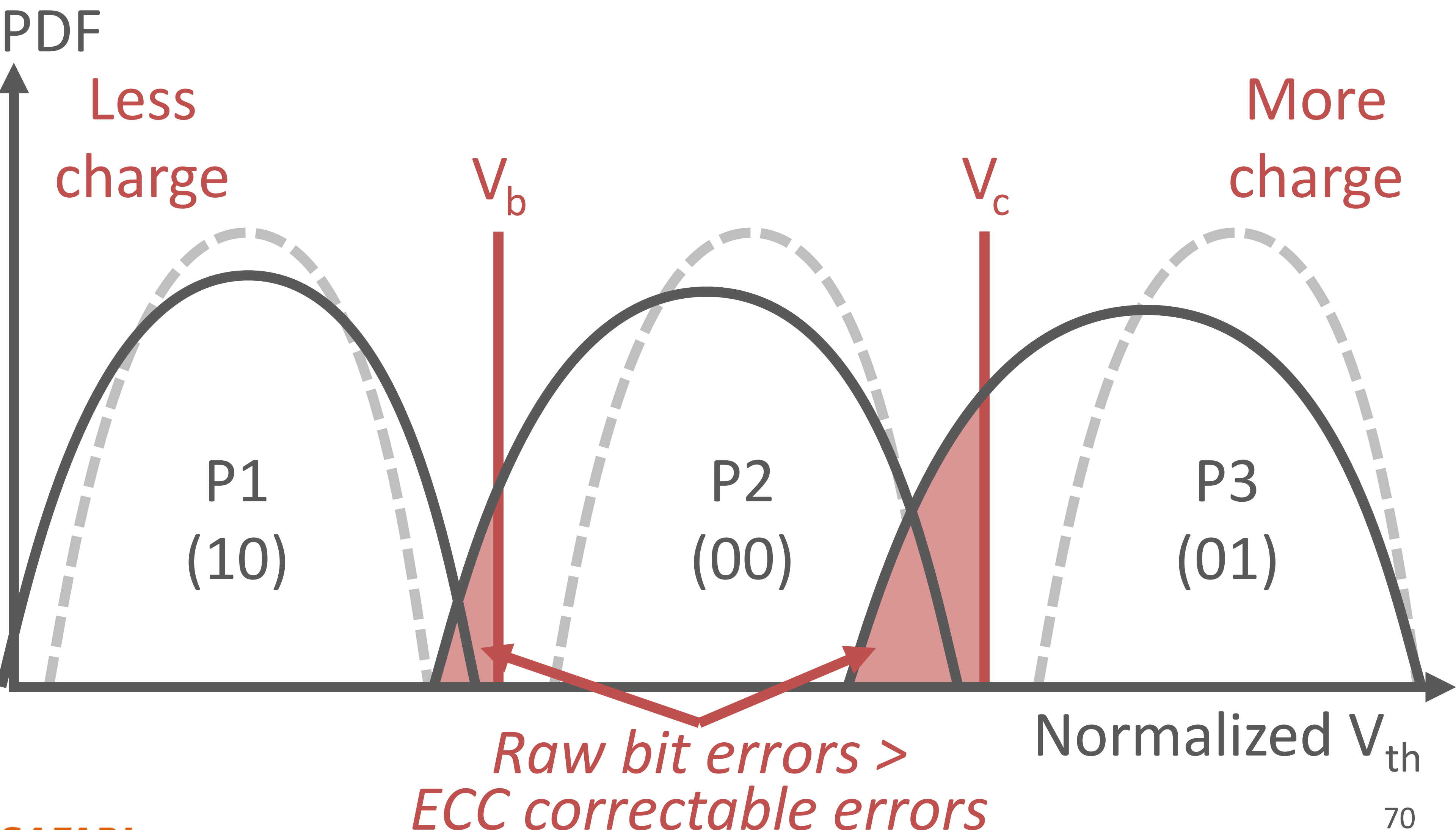
Threshold voltage reduces over time

Old data



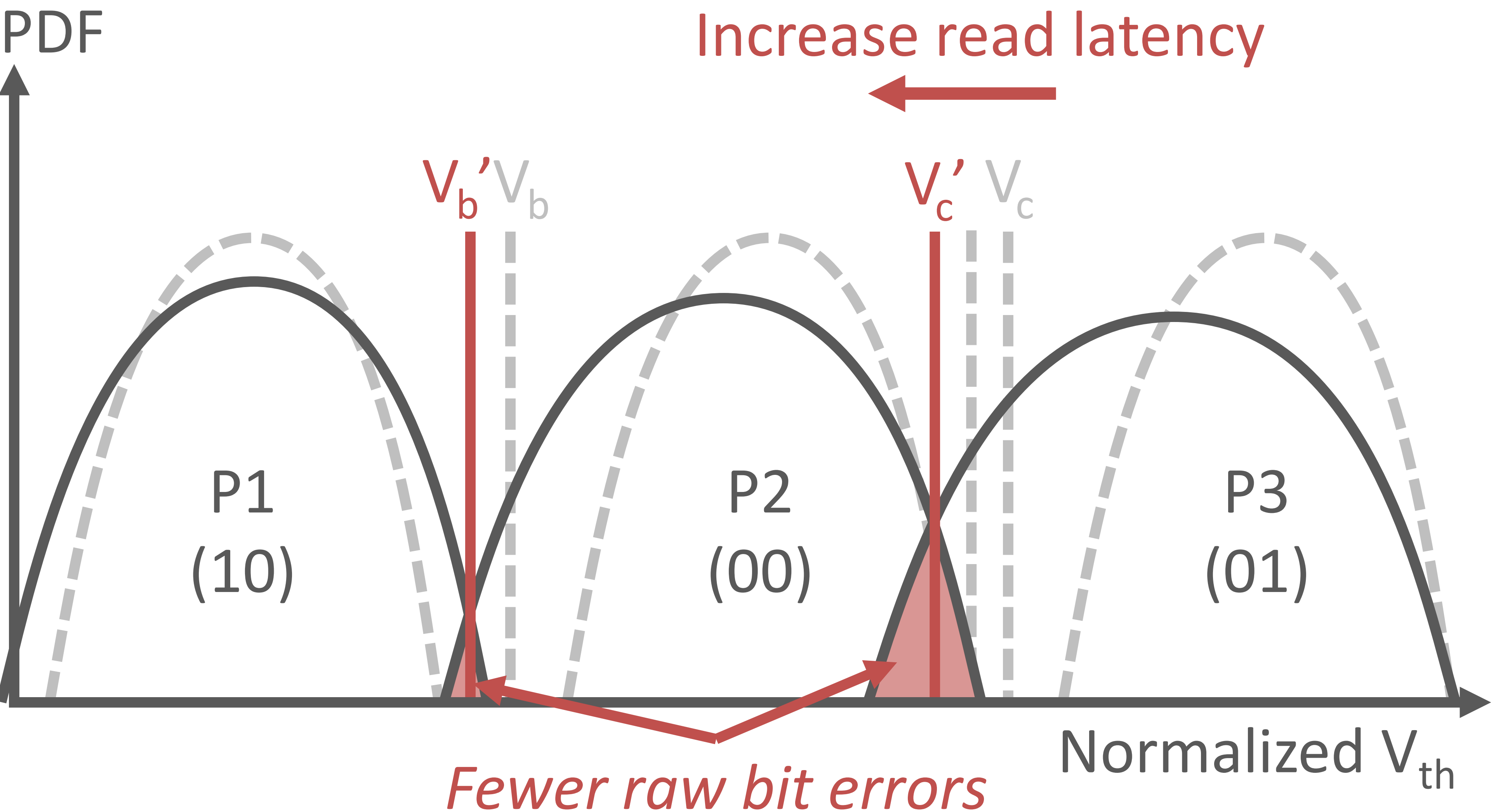
First read attempt fails

Old data



Read-retry

Old data



Why is old data slower?

Retention loss

→ *Leak charge over time*

→ *Generate retention errors*

→ *Require read-retry*

→ *Longer read latency*

Characterize retention loss in real NAND chip

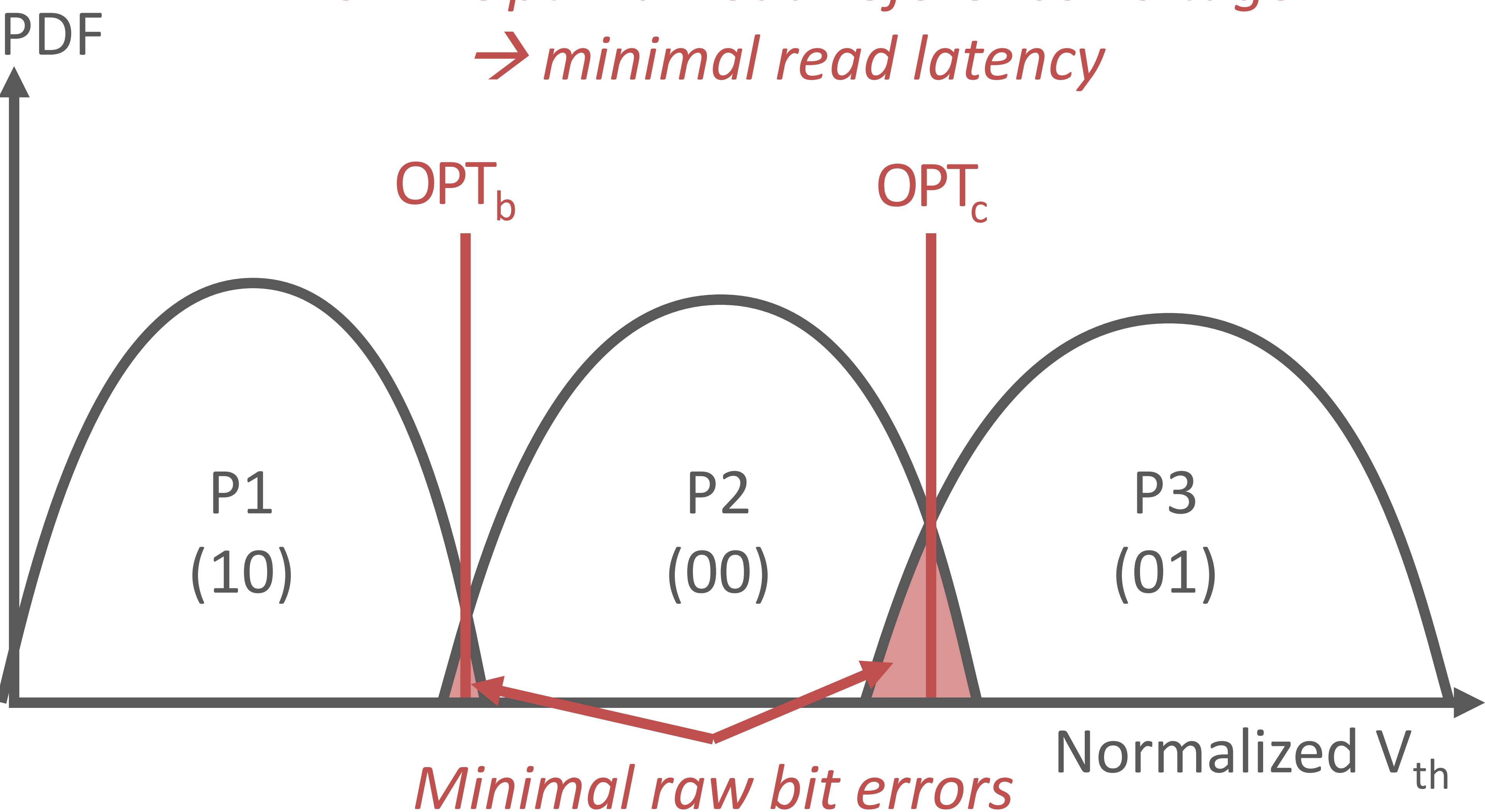
Optimize read performance for old data

Recover old data after failure

The ideal read voltage

Old data

OPT: Optimal read reference voltage
→ minimal read latency



In reality

- *OPT changes over time due to retention loss*

- *Luckily, OPT change is:*
  - Gradual
  - Uni-directional (decreases over time)

Retention Optimized Reading (ROR)

Components:

1. Online pre-optimization algorithm

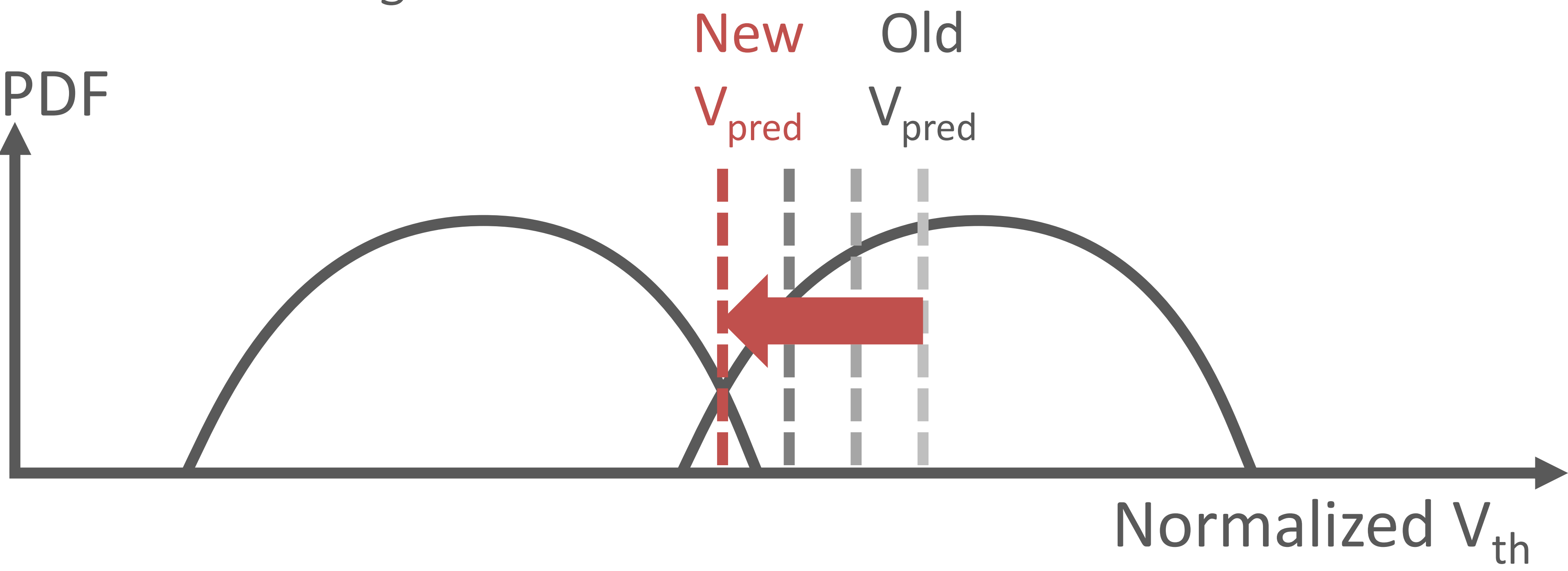
- Learns and records OPT
- Performs in the background once every day

2. Simpler read-retry technique

- If recorded OPT is out-of-date, read-retry with *lower* voltage

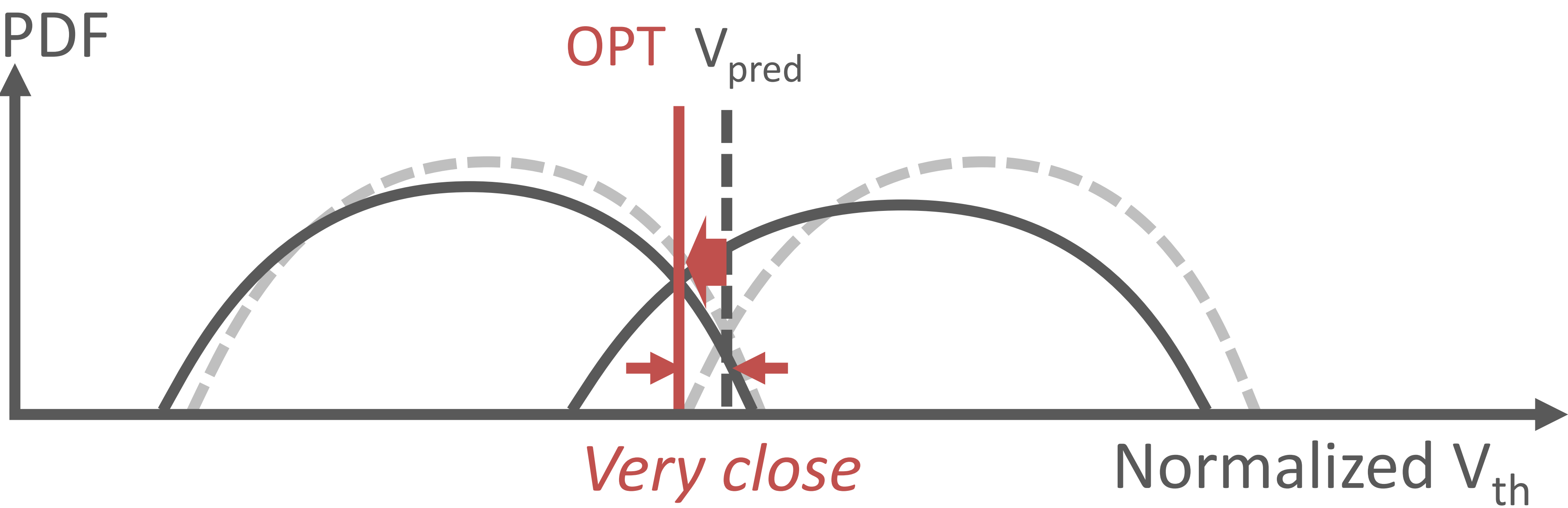
1. Online Pre-Optimization Algorithm

- *Triggered periodically (e.g., per day)*
- *Find and record an OPT as per-block V_{pred}*
- *Performed in background*
- *Small storage overhead*



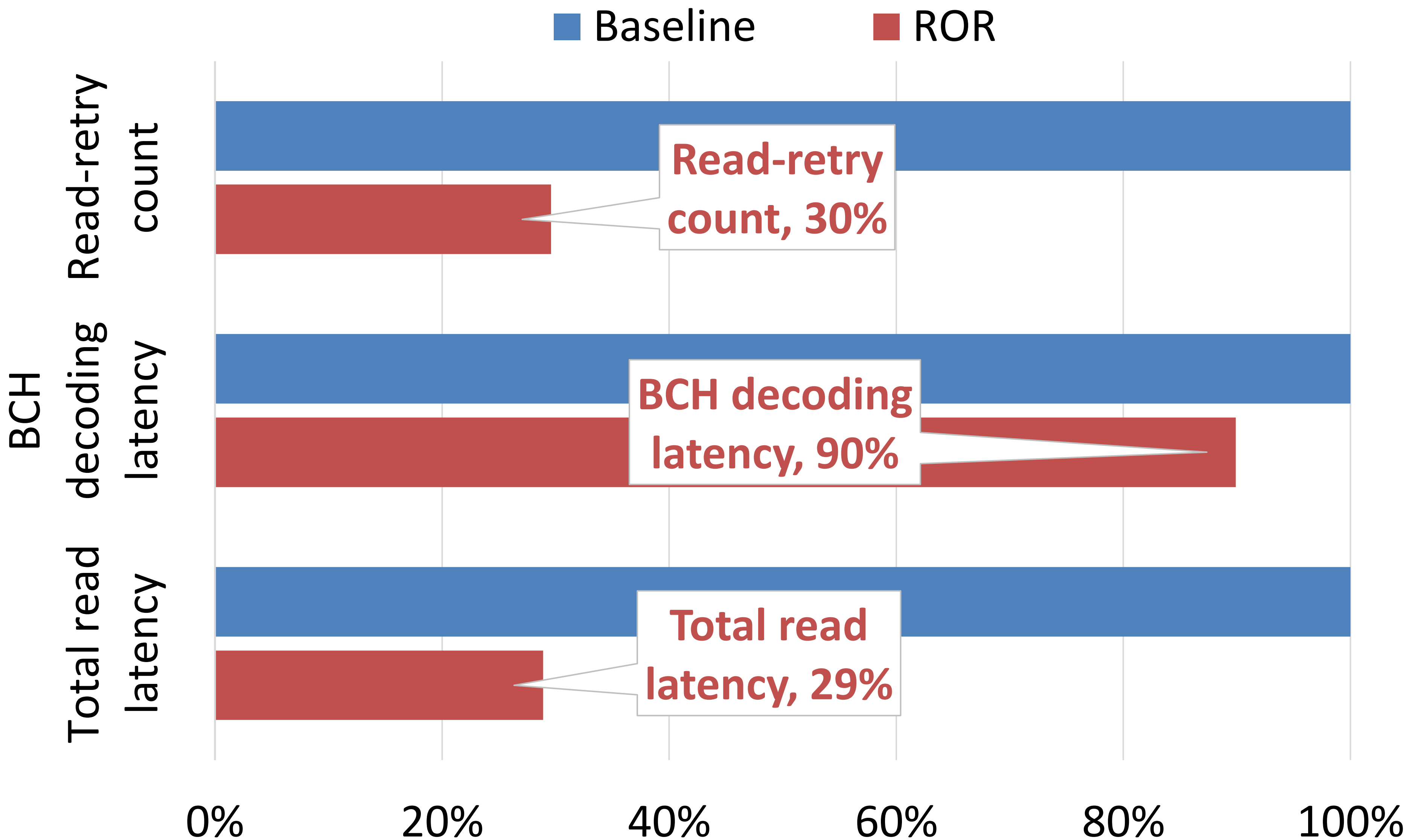
2. Improved Read-Retry Technique

- *Performed as normal read*
- *V_{pred} already close to actual OPT*
- *Decrease V_{ref} if V_{pred} fails, and retry*



SAFARI

ROR result



Baseline
ROR
Read-retry count
BCH decoding latency
Total read latency
Read-retry count, 30%
BCH decoding latency, 90%
Total read latency, 29%
0%
20%
40%
60%
80%
100%

Retention optimized reading

Retention loss → longer read latency

Optimal read reference voltage (OPT)

→ Shortest read latency

→ Decreases gradually over time (retention)

→ Learn OPT periodically

→ Minimize read-retry & RBER

→ Shorter read latency

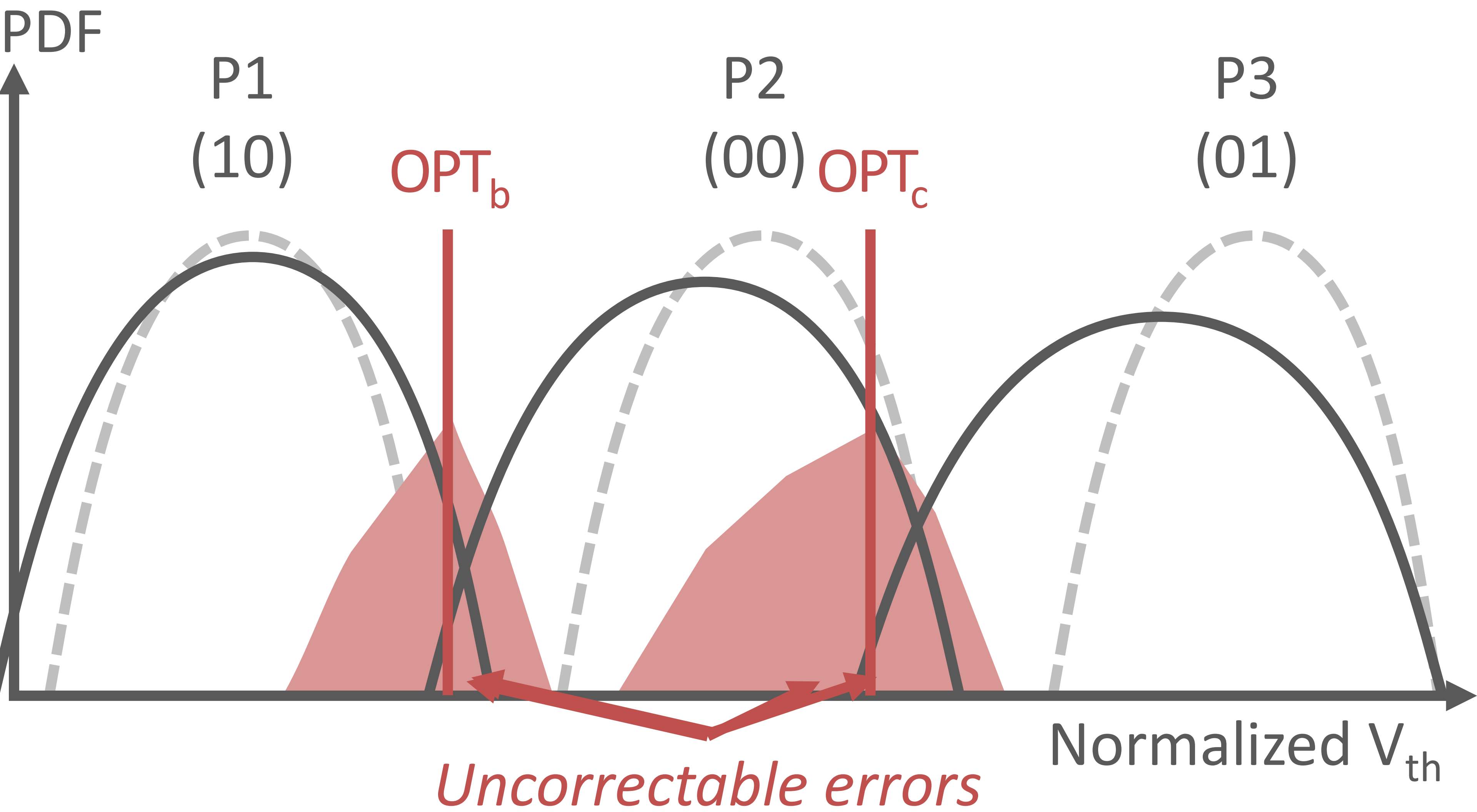
Characterize retention loss in real NAND chip

Optimize read performance for old data

Recover old data after failure

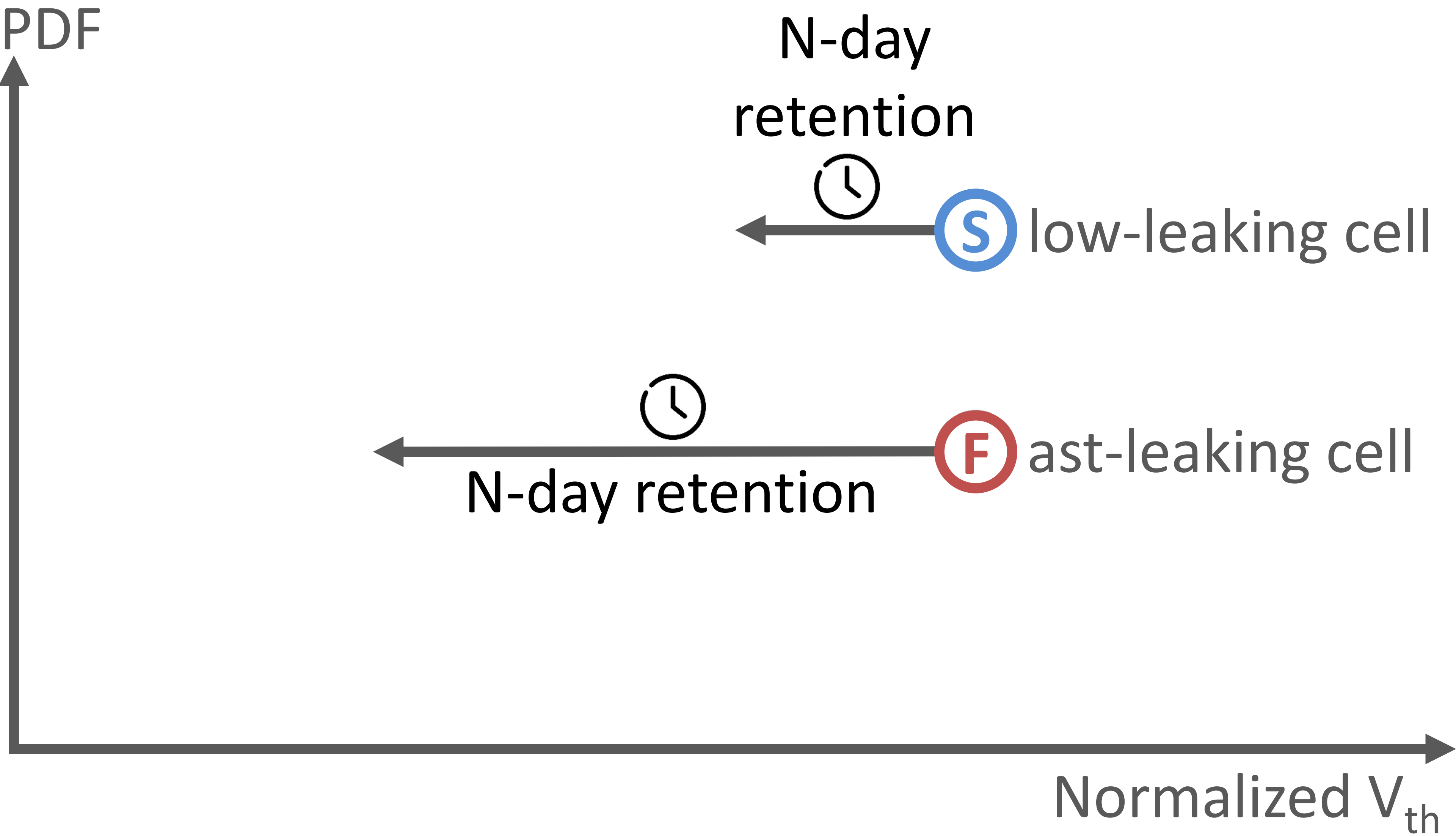
Retention failure

Very old data



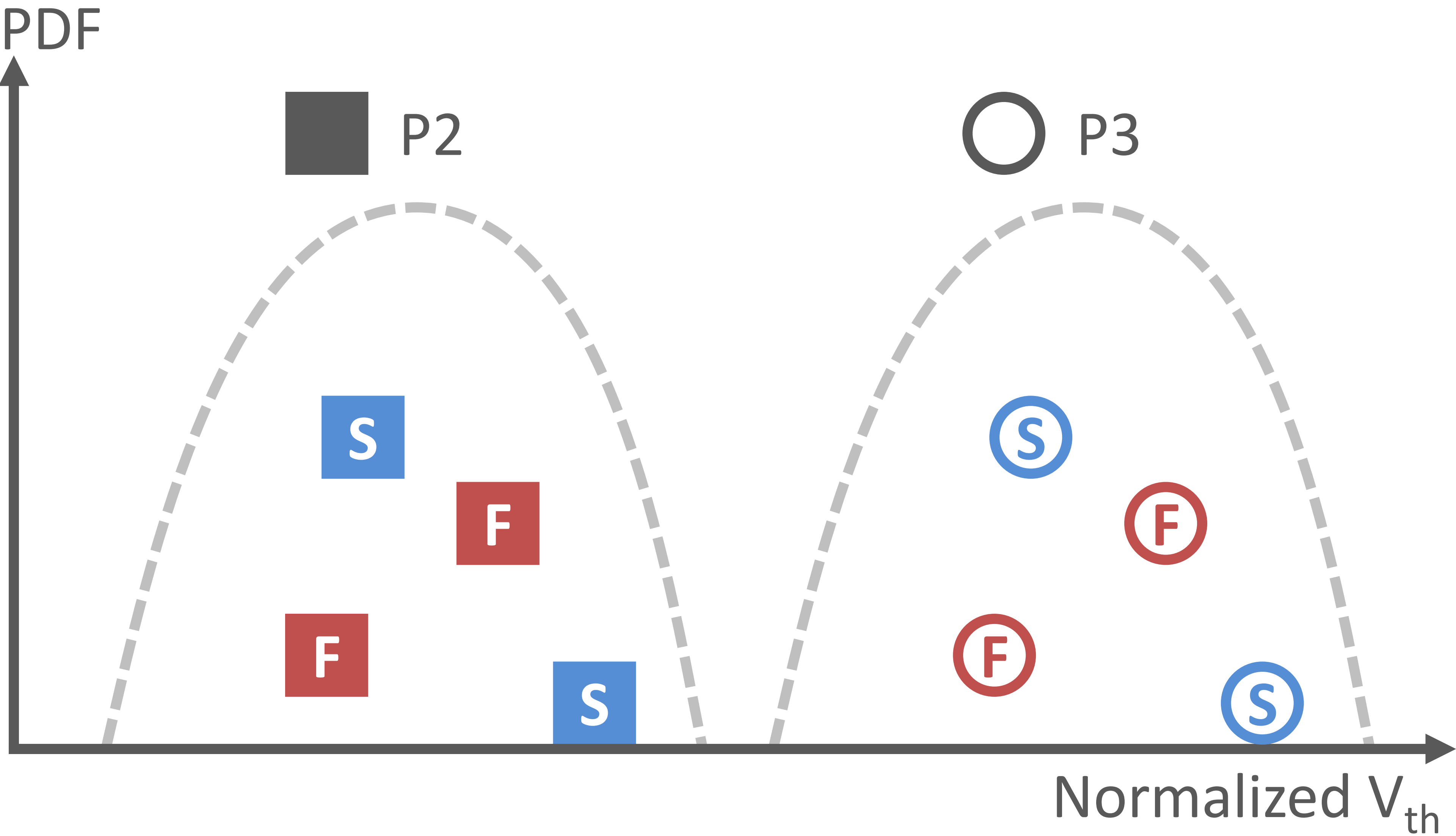
SAFARI

Leakage speed variation



SAFARI

A simplified example



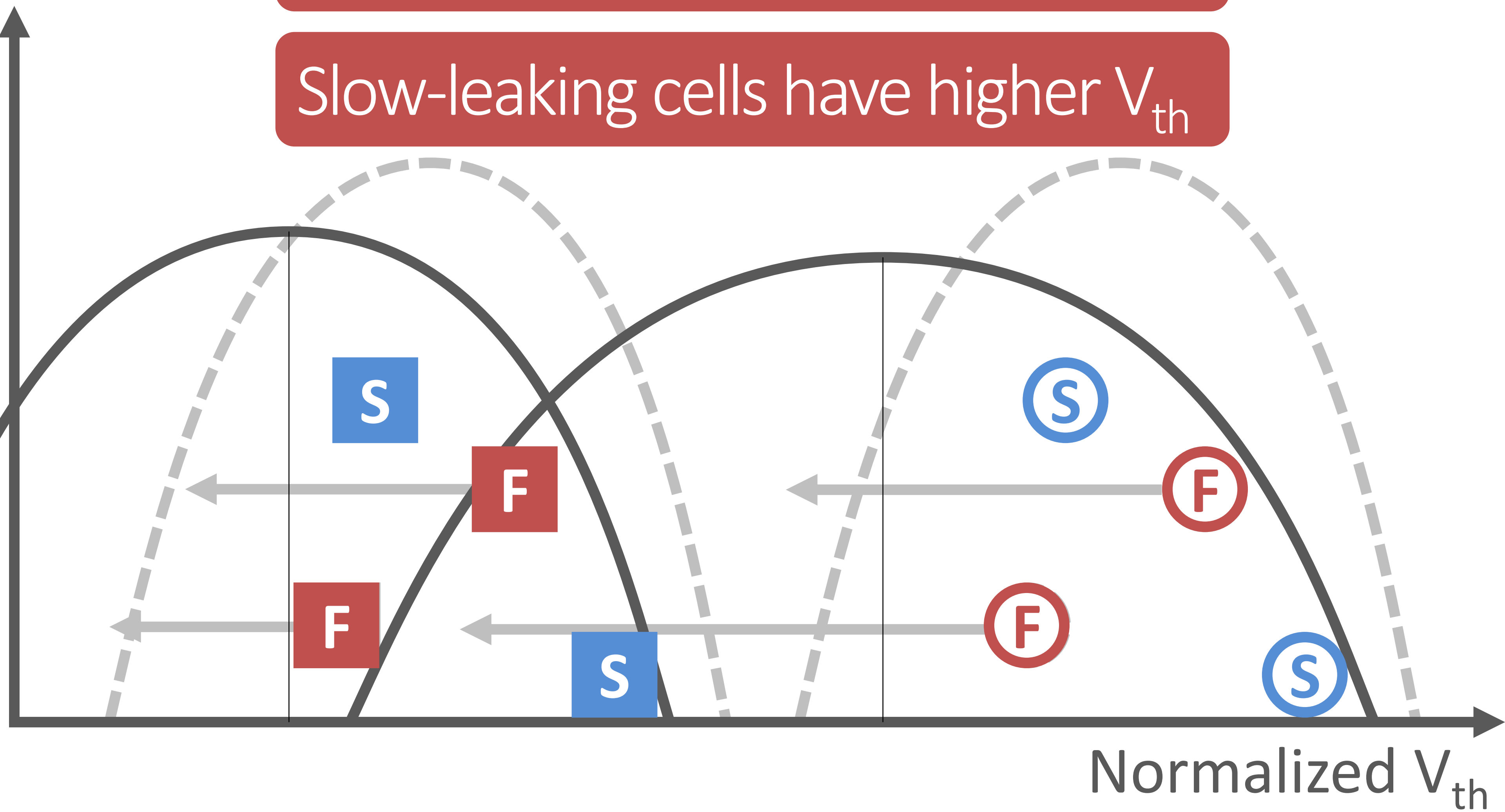
Reading very old data

Very old

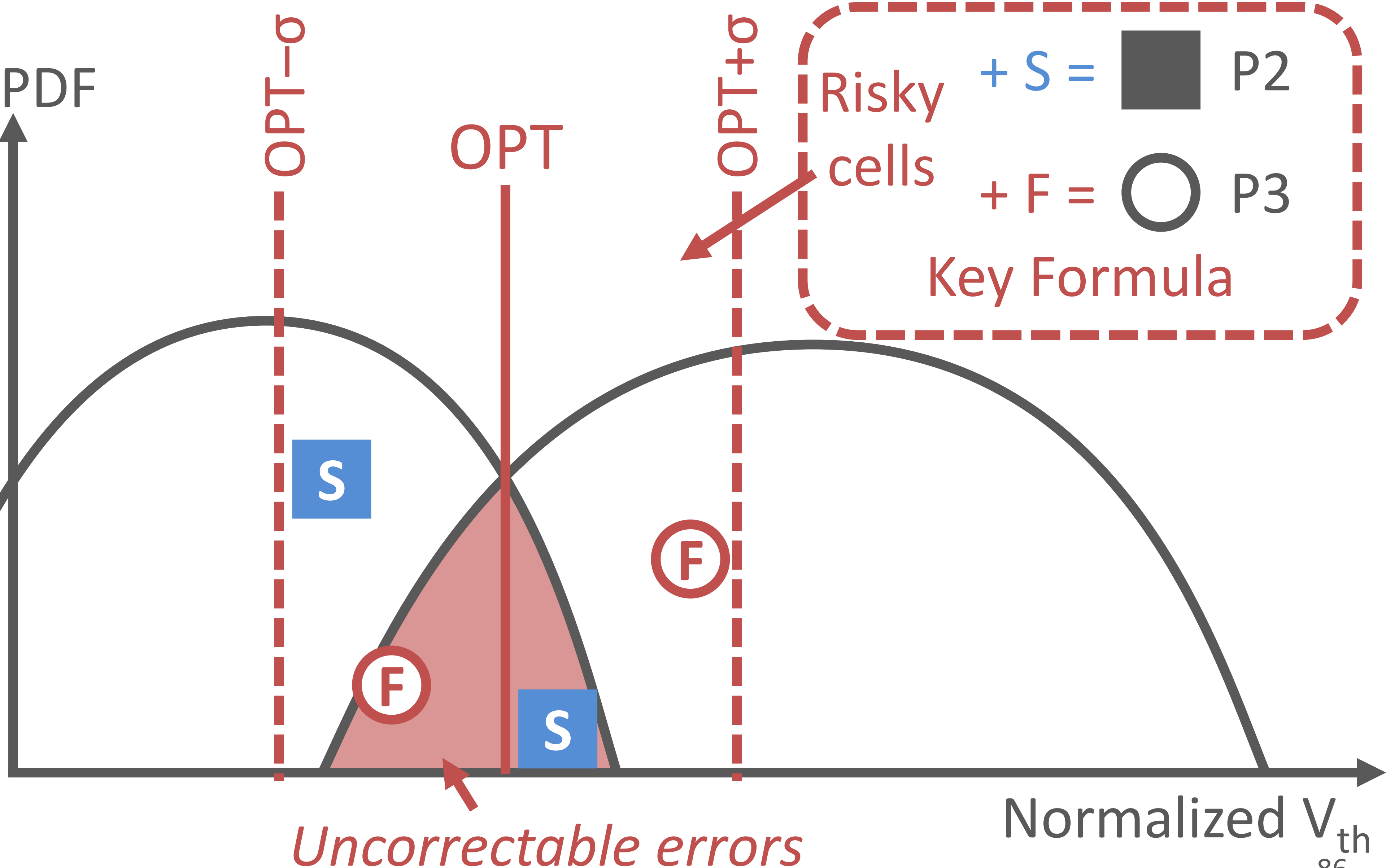
Fast-leaking cells have lower V_{th}

Slow-leaking cells have higher V_{th}

PDF

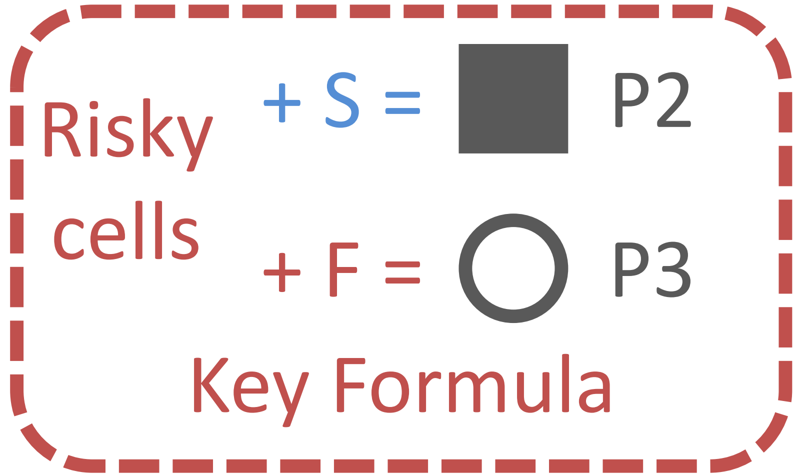


“Risky” cells


PDF
OPT−σ
OPT
OPT+σ
Risky cells
+ S =
P2
+ F =
P3
Key Formula
S
F
F
S
Uncorrectable errors
Normalized V_{th}

Retention Failure Recovery (RFR)

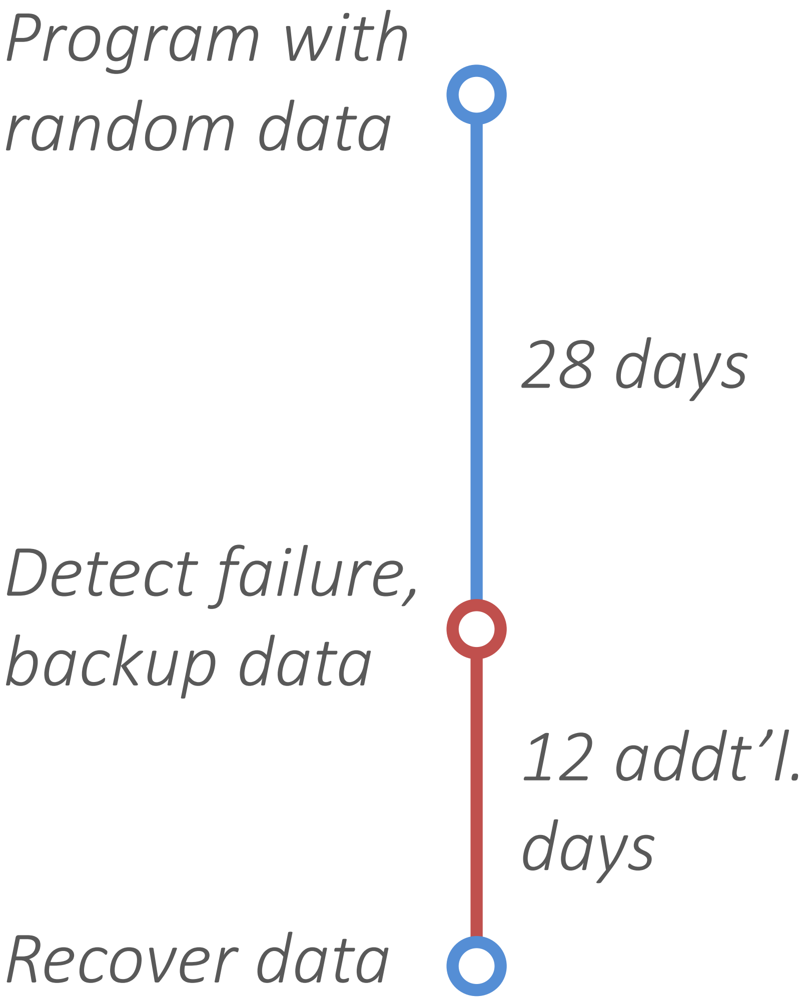
Key idea: *Guess original state of the cell from its leakage speed property*



Three steps

1. *Identify risky cells*
2. *Identify fast-/slow-leaking cells*
3. *Guess original states*

RFR Evaluation



- *Expect to eliminate 50% of raw bit errors*
- *ECC can correct remaining errors*

Characterize retention loss in real NAND chip

Optimize read performance for old data

Recover old data after failure

Conclusion

Retention loss → Longer read latency

Retention optimized reading (ROR)

→ Learns OPT periodically

→ 71% shorter read latency

Retention failure recovery (RFR)

→ Use leakage property to guess correct state

→ 50% error reduction before ECC correction

→ Recover data after failure

More on Flash Read Disturb Errors

- Yu Cai, Yixin Luo, Erich F. Haratsch, Ken Mai, and Onur Mutlu, **"Data Retention in MLC NAND Flash Memory: Characterization, Optimization and Recovery"** *Proceedings of the 21st International Symposium on High-Performance Computer Architecture (**HPCA**)*, Bay Area, CA, February 2015. [Slides (pptx) (pdf)]

Data Retention in MLC NAND Flash Memory: Characterization, Optimization, and Recovery

Yu Cai, Yixin Luo, Erich F. Haratsch*, Ken Mai, Onur Mutlu
Carnegie Mellon University, *LSI Corporation
yucaicai@gmail.com, yixinluo@cs.cmu.edu, erich.haratsch@lsi.com, {kenmai, omutlu}@ece.cmu.edu

Agenda

- Background, Motivation and Approach
- Experimental Characterization Methodology
- Error Analysis and Management
  - Main Characterization Results
  - Retention-Aware Error Management
  - Threshold Voltage and Program Interference Analysis
  - Read Reference Voltage Prediction
  - Neighbor-Assisted Error Correction
  - Read Disturb Error Handling
  - Retention Error Handling
  - Large Scale Field Analysis
  - 3D NAND Flash Memory Reliability
- Summary

Large Scale Field Analysis of Flash Memory Errors

SSD Error Analysis of Facebook Systems

- First large-scale field study of flash memory errors
- Justin Meza, Qiang Wu, Sanjeev Kumar, and Onur Mutlu,
  "A Large-Scale Study of Flash Memory Errors in the Field"
  *Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (**SIGMETRICS**)*, Portland, OR, June 2015.
  [Slides (pptx) (pdf)] [Coverage at ZDNet] [Coverage on The Register] [Coverage on TechSpot] [Coverage on The Tech Report]

A Large-Scale Study of Flash Memory Failures in the Field

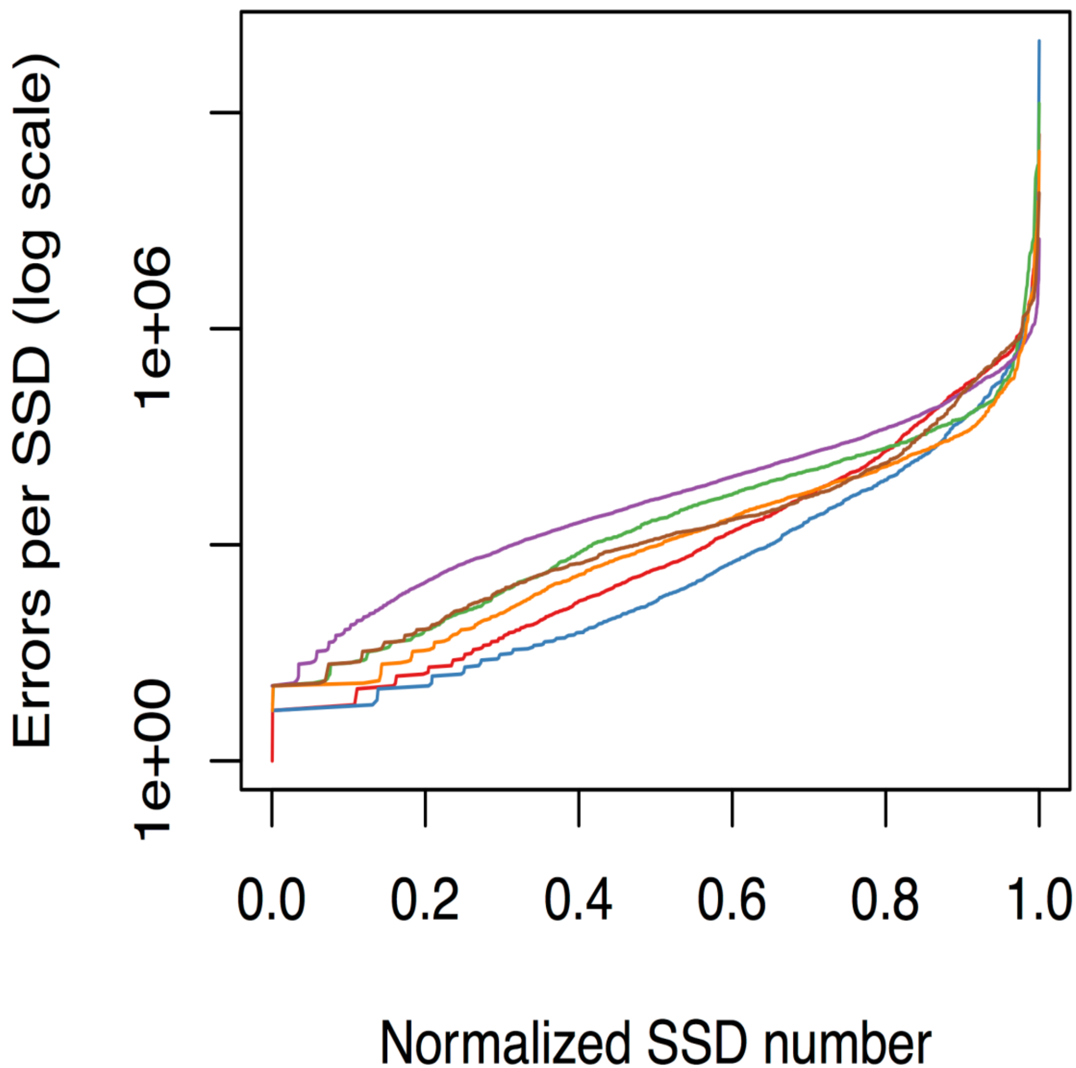
Justin Meza
Carnegie Mellon University
meza@cmu.edu

Qiang Wu
Facebook, Inc.
qwu@fb.com

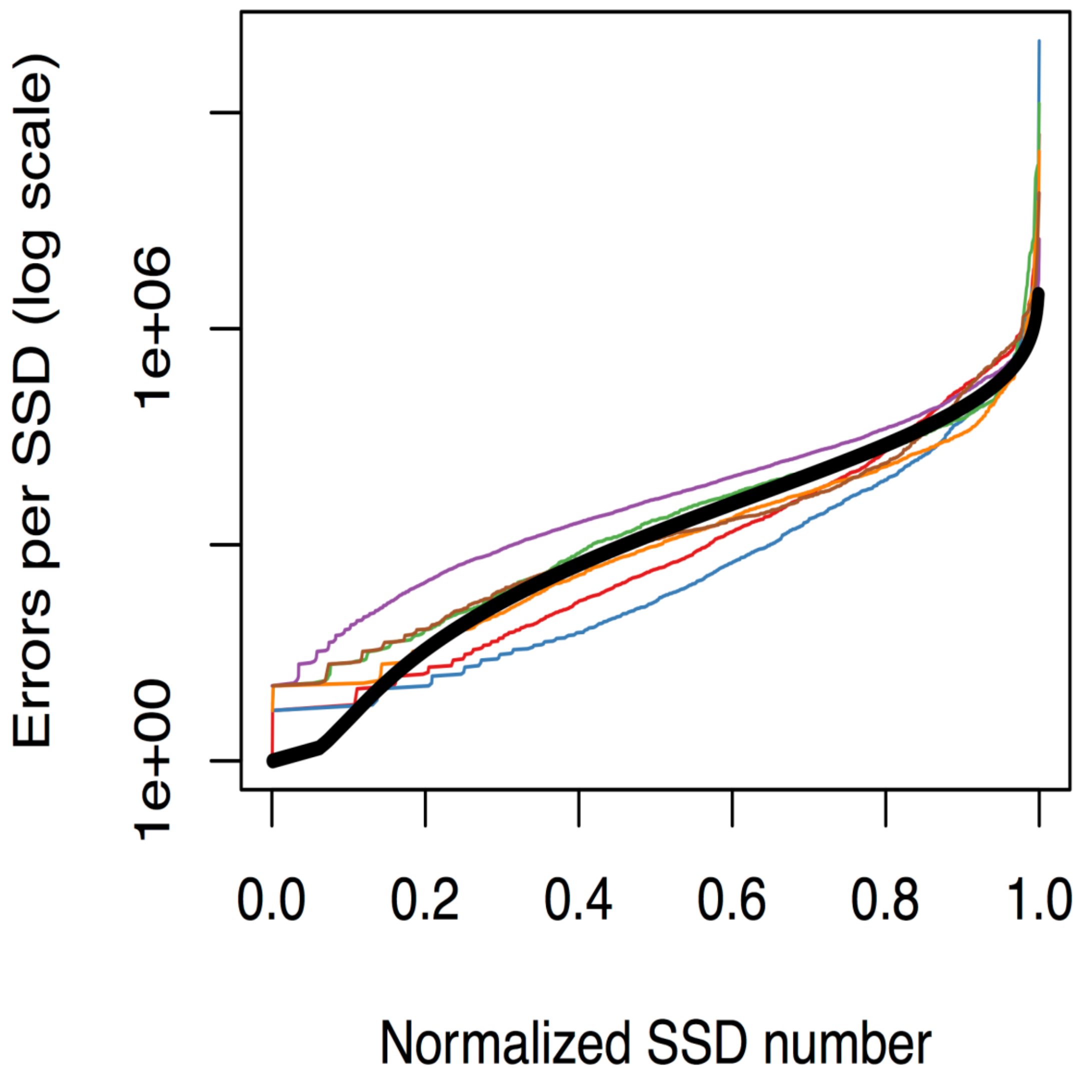
Sanjeev Kumar
Facebook, Inc.
skumar@fb.com

Onur Mutlu
Carnegie Mellon University
onur@cmu.edu

A few SSDs cause most errors



A few SSDs cause most errors



Summary

SSD lifecycle

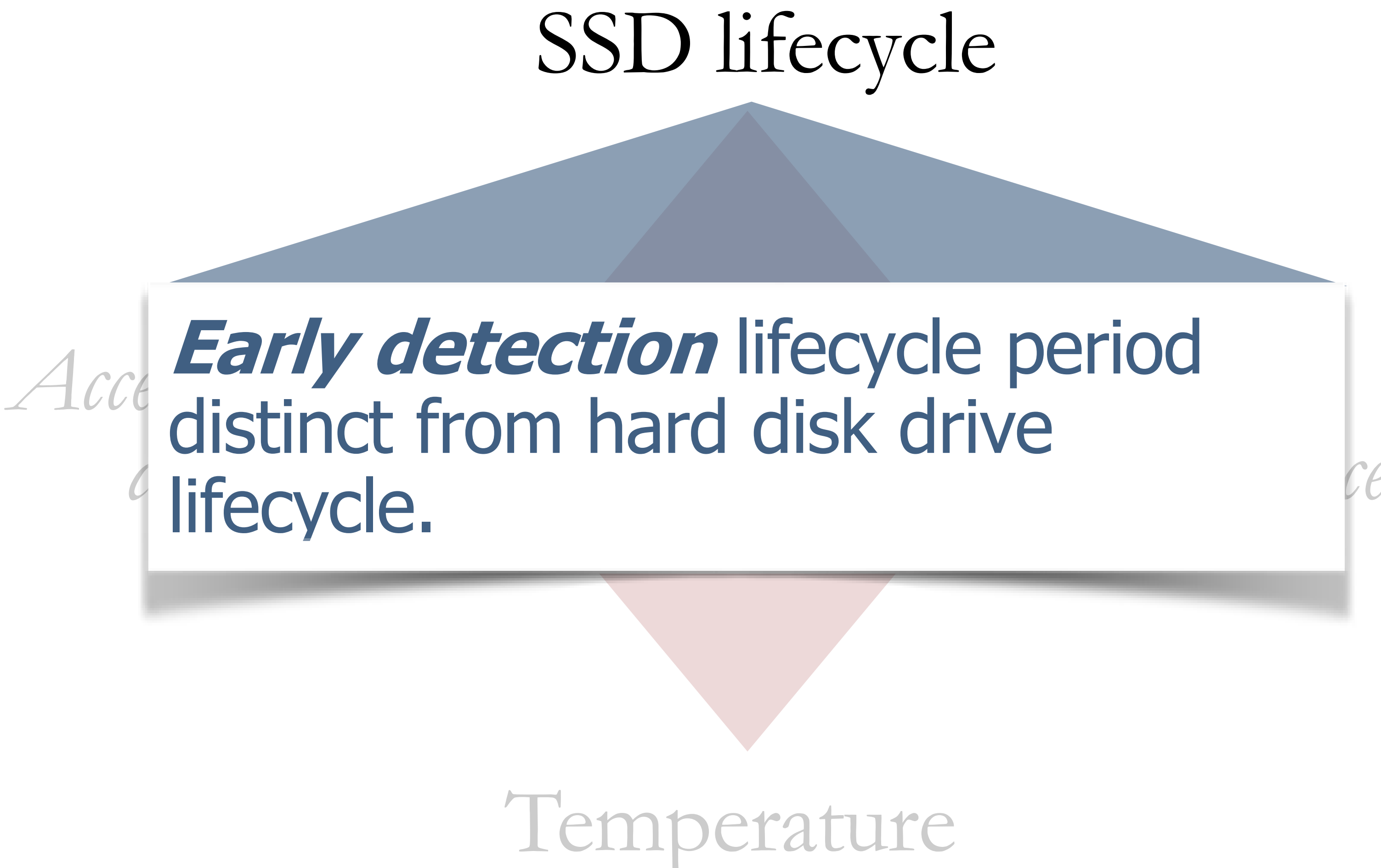
Access pattern dependence

New reliability trends

Read disturbance

Temperature

Summary


SSD lifecycle
Early detection lifecycle period distinct from hard disk drive lifecycle.
Temperature

SSD lifecycle

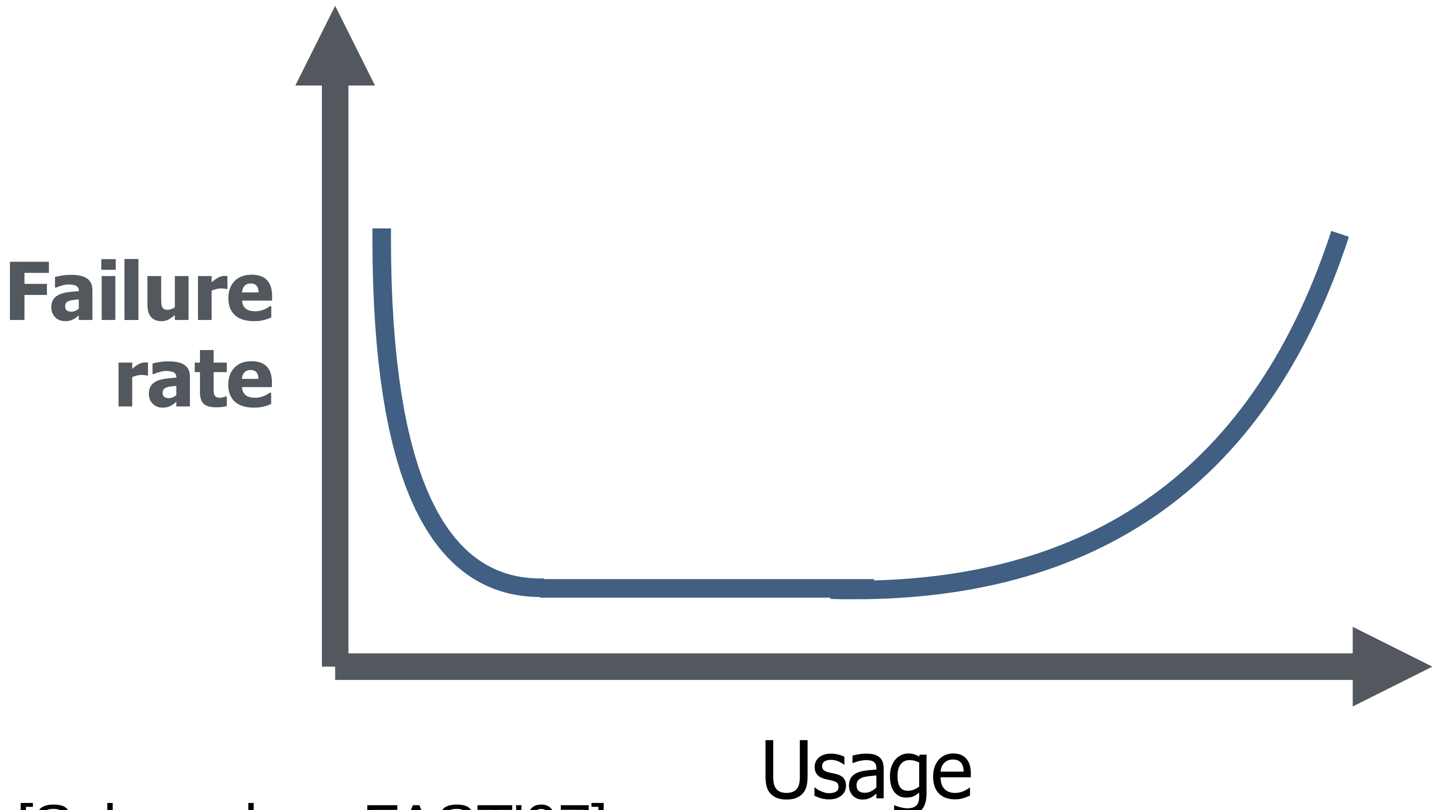
Access pattern dependence

New reliability trends

Read disturbance

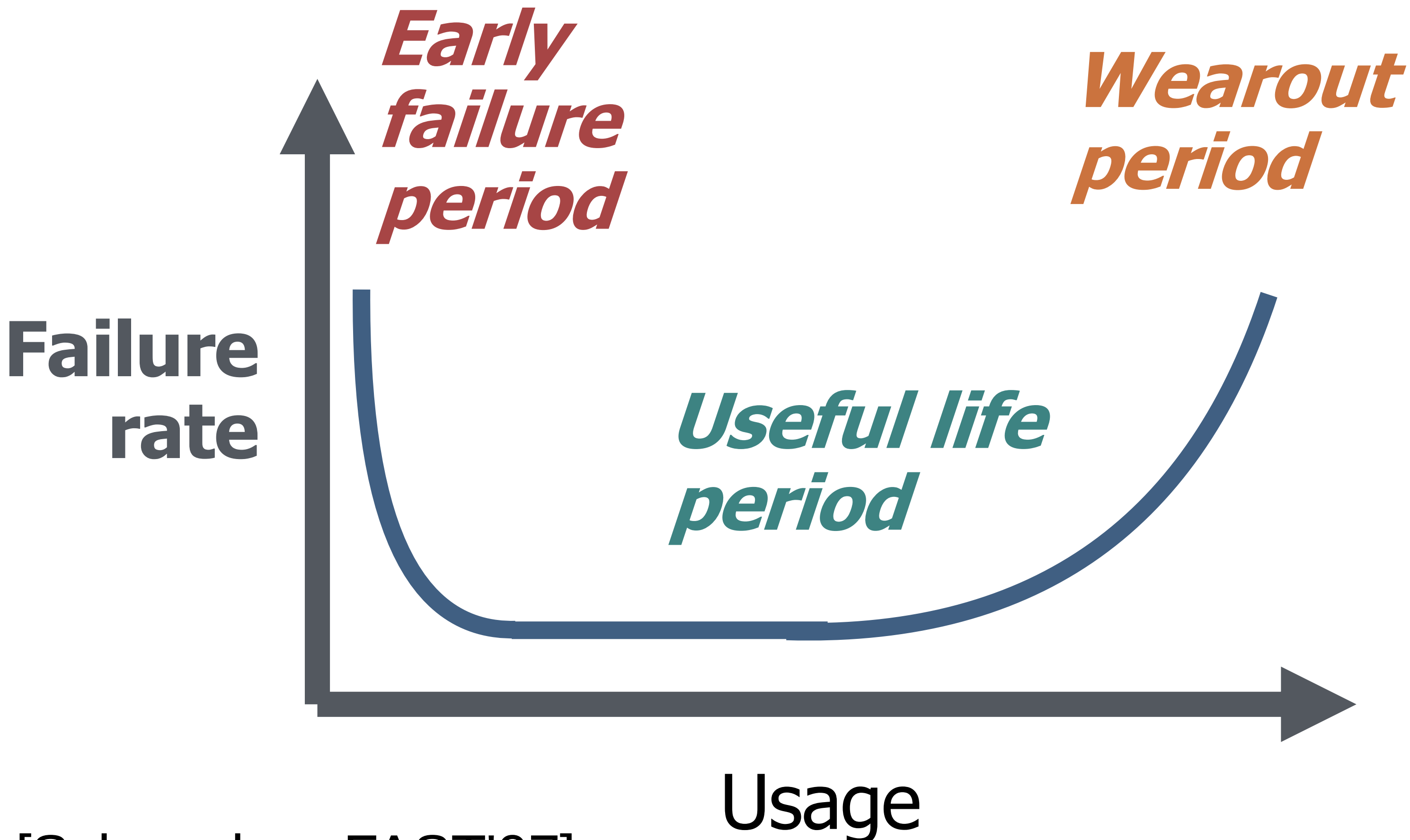
Temperature

Storage lifecycle background: the bathtub curve for disk drives



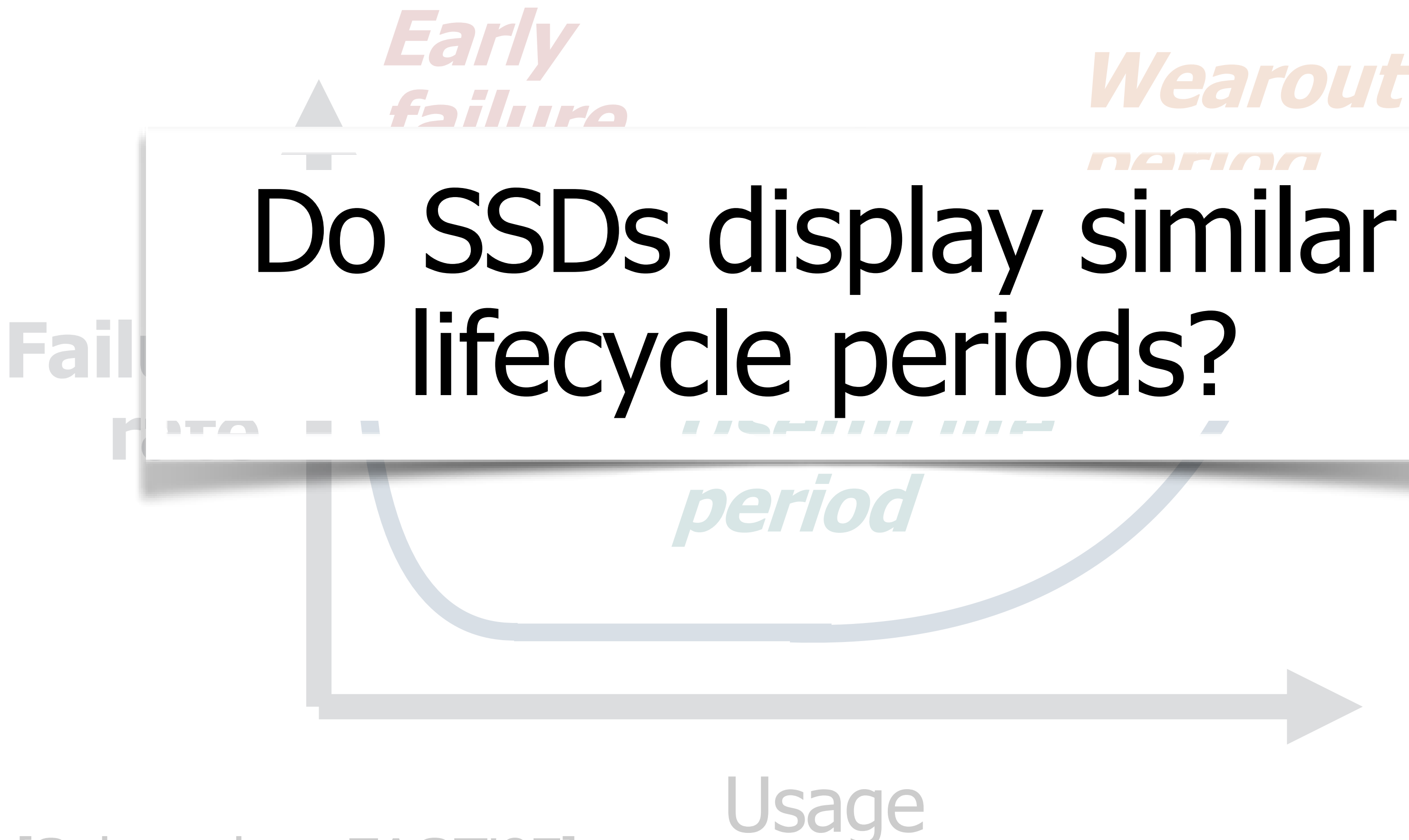
[Schroeder+,FAST'07]

Storage lifecycle background: the bathtub curve for disk drives



[Schroeder+,FAST'07]

Storage lifecycle background: the bathtub curve for disk drives

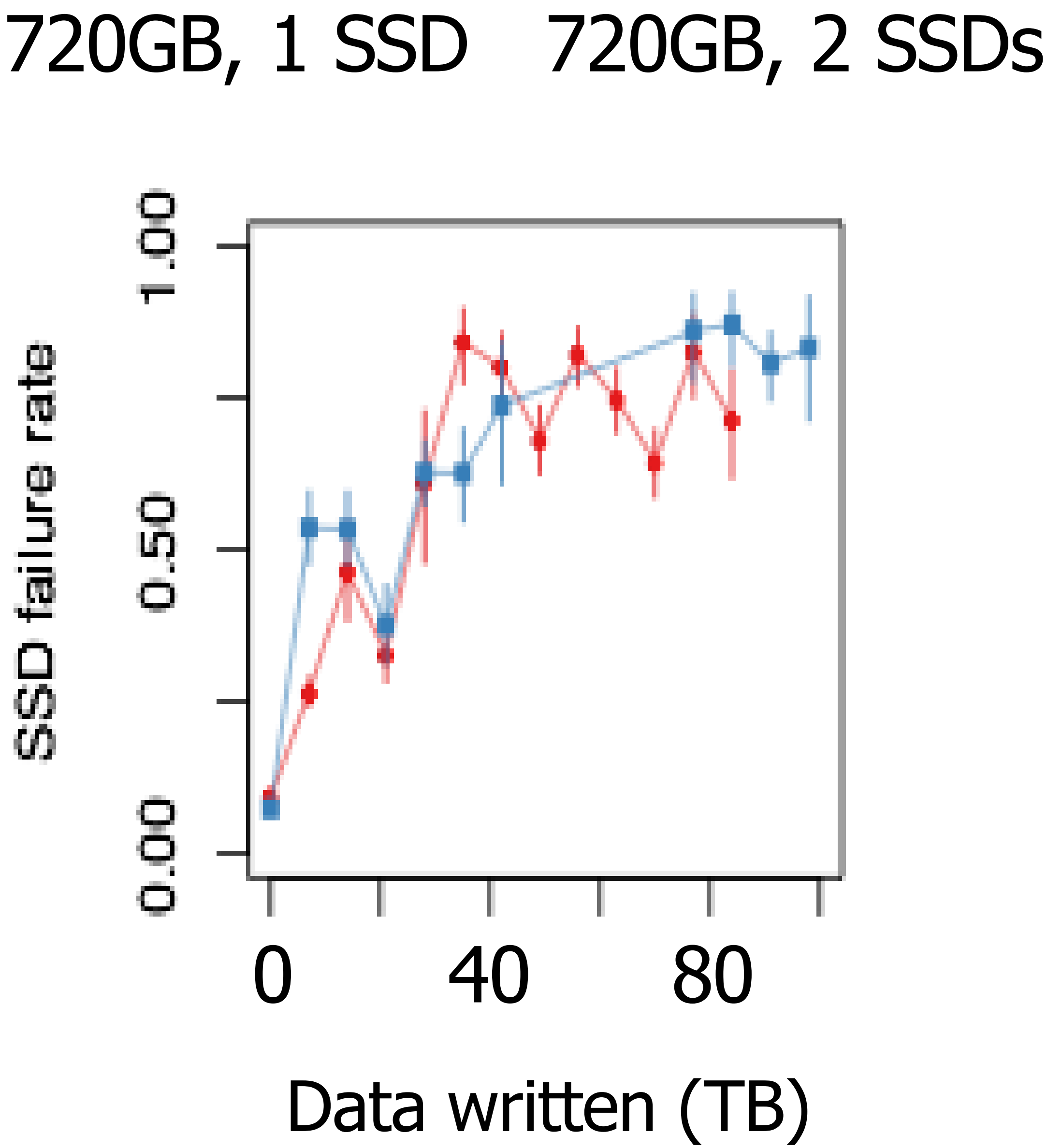


Do SSDs display similar lifecycle periods?

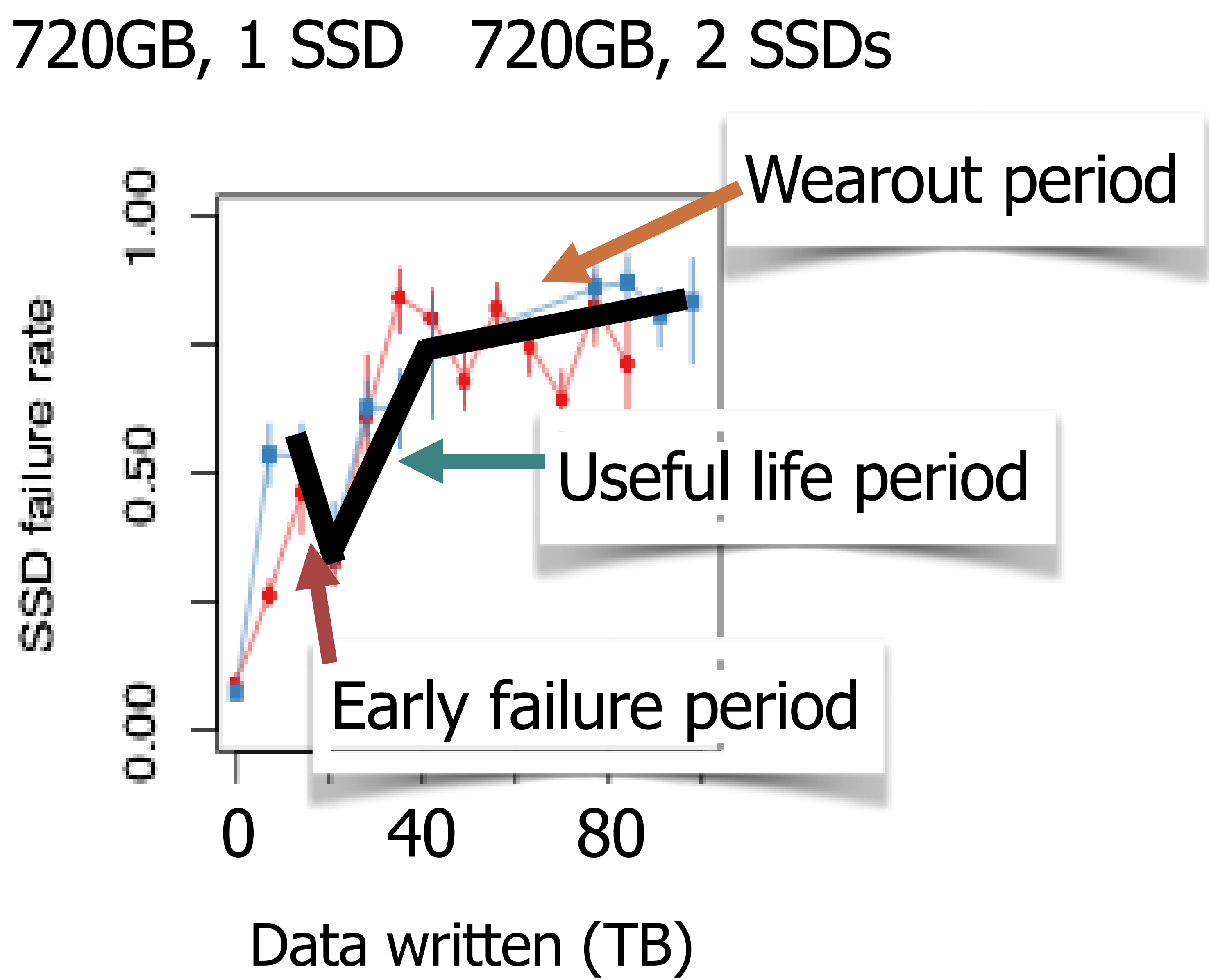
[Schroeder+,FAST'07]

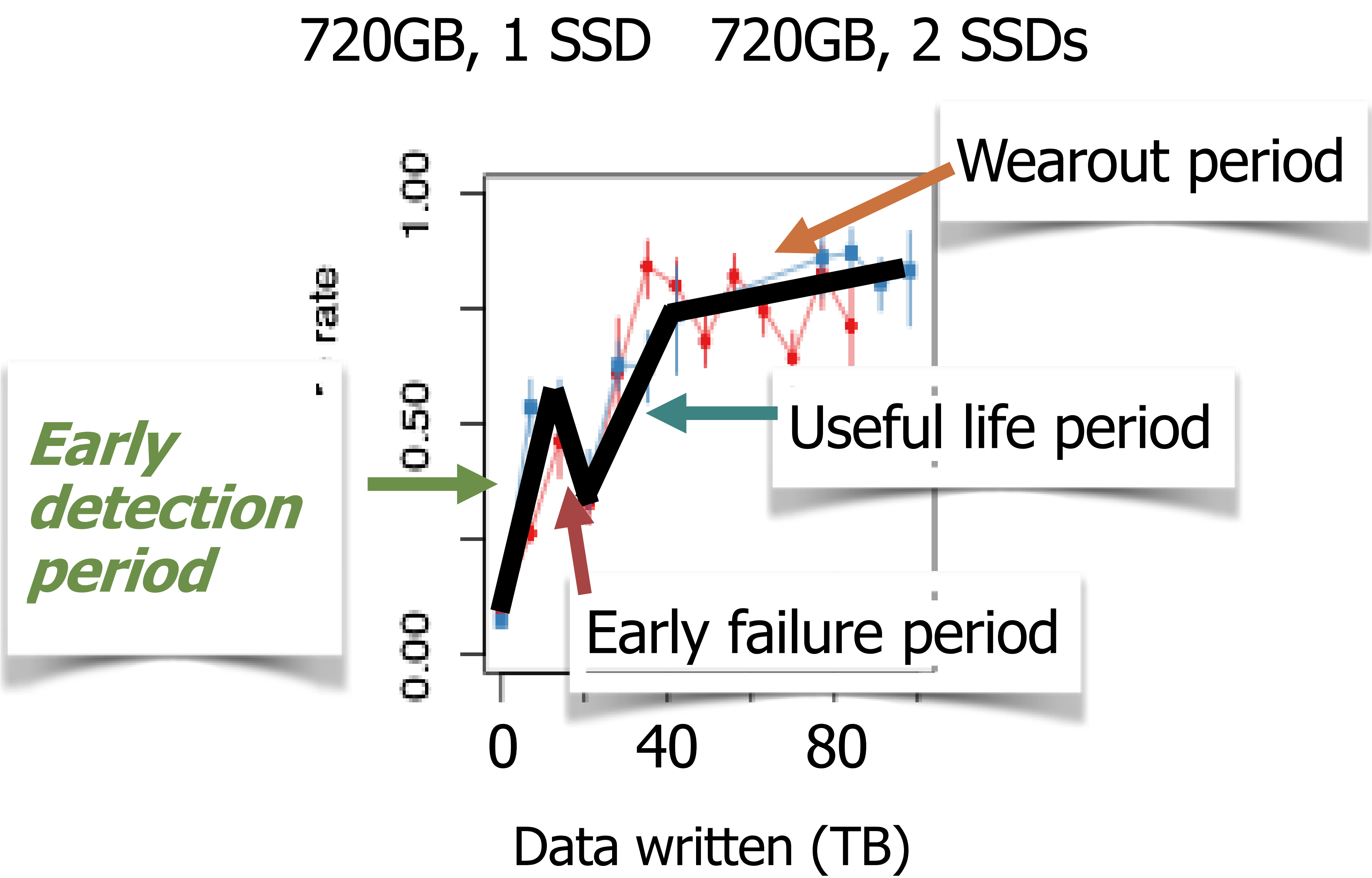
Use data written to flash
to examine SSD lifecycle

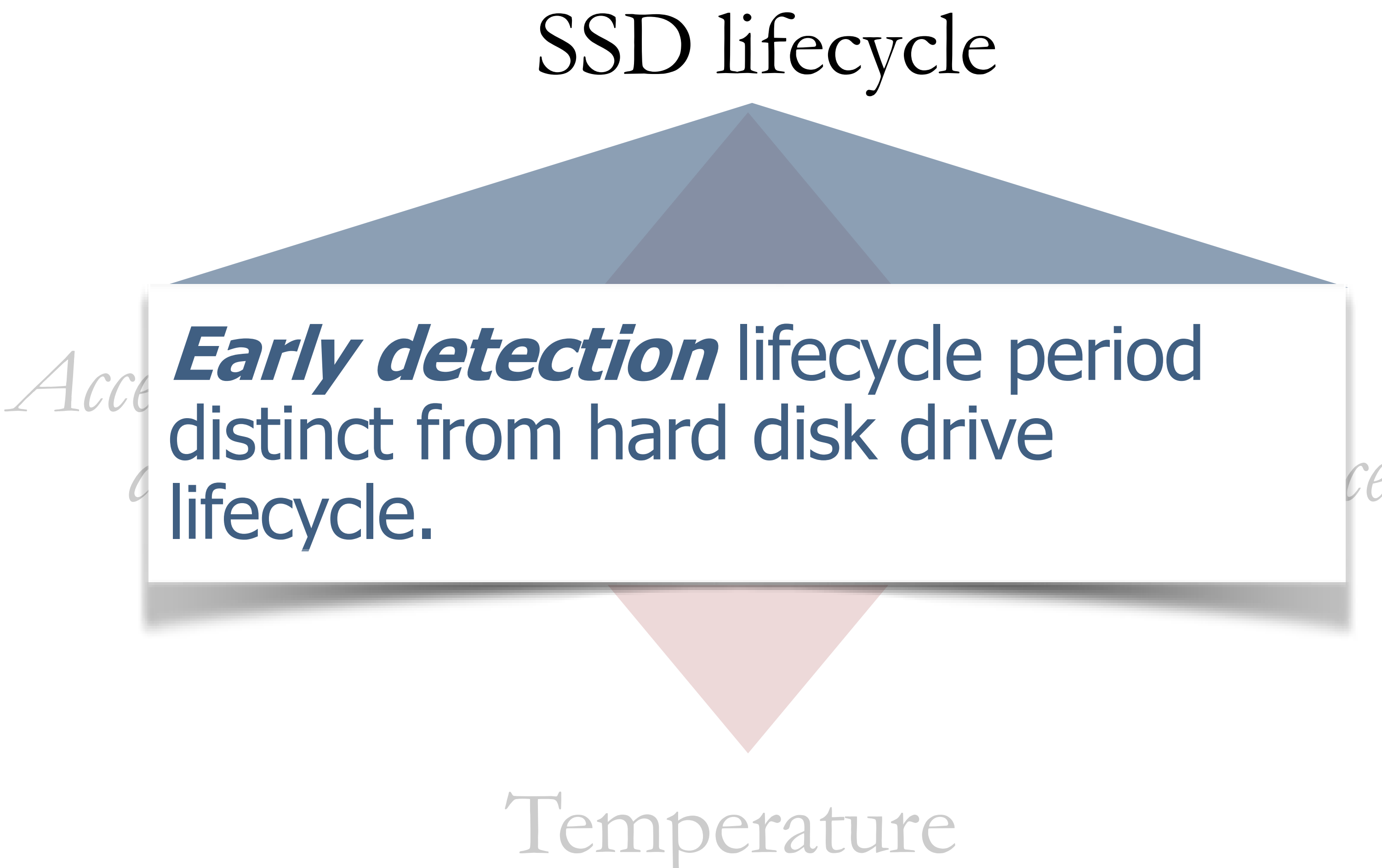
(time-independent utilization metric)



720GB, 1 SSD
720GB, 2 SSDs
SSD failure rate
0.00
0.50
1.00
0
40
80
Data written (TB)


720GB, 1 SSD
720GB, 2 SSDs
Wearout period
Useful life period
Early failure period
SSD failure rate
1.00
0.50
0.00
0
40
80
Data written (TB)


720GB, 1 SSD
720GB, 2 SSDs
Wearout period
Useful life period
Early detection period
Early failure period
1.00
0.50
0.00
rate
0
40
80
Data written (TB)


SSD lifecycle
Early detection lifecycle period distinct from hard disk drive lifecycle.
Temperature

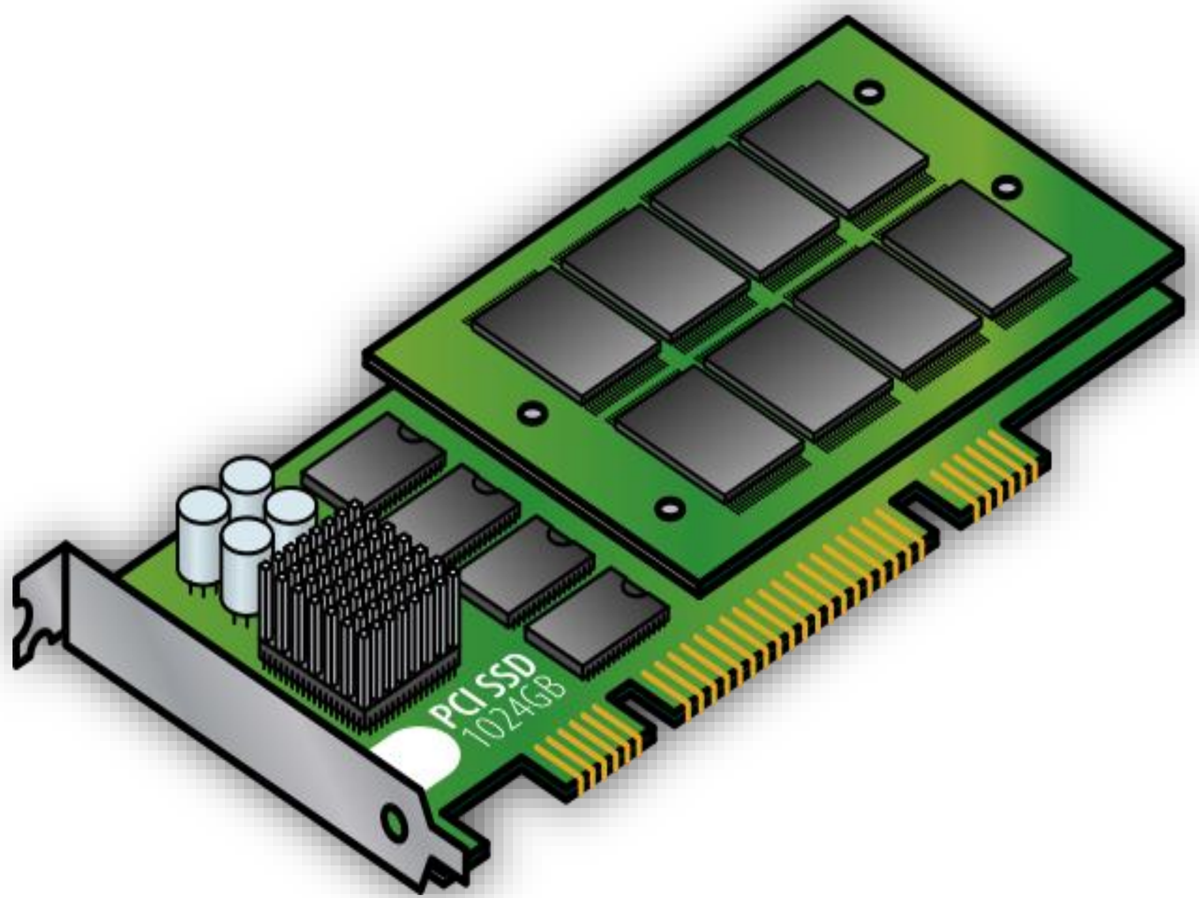
SSD lifecycle

Access pattern dependence

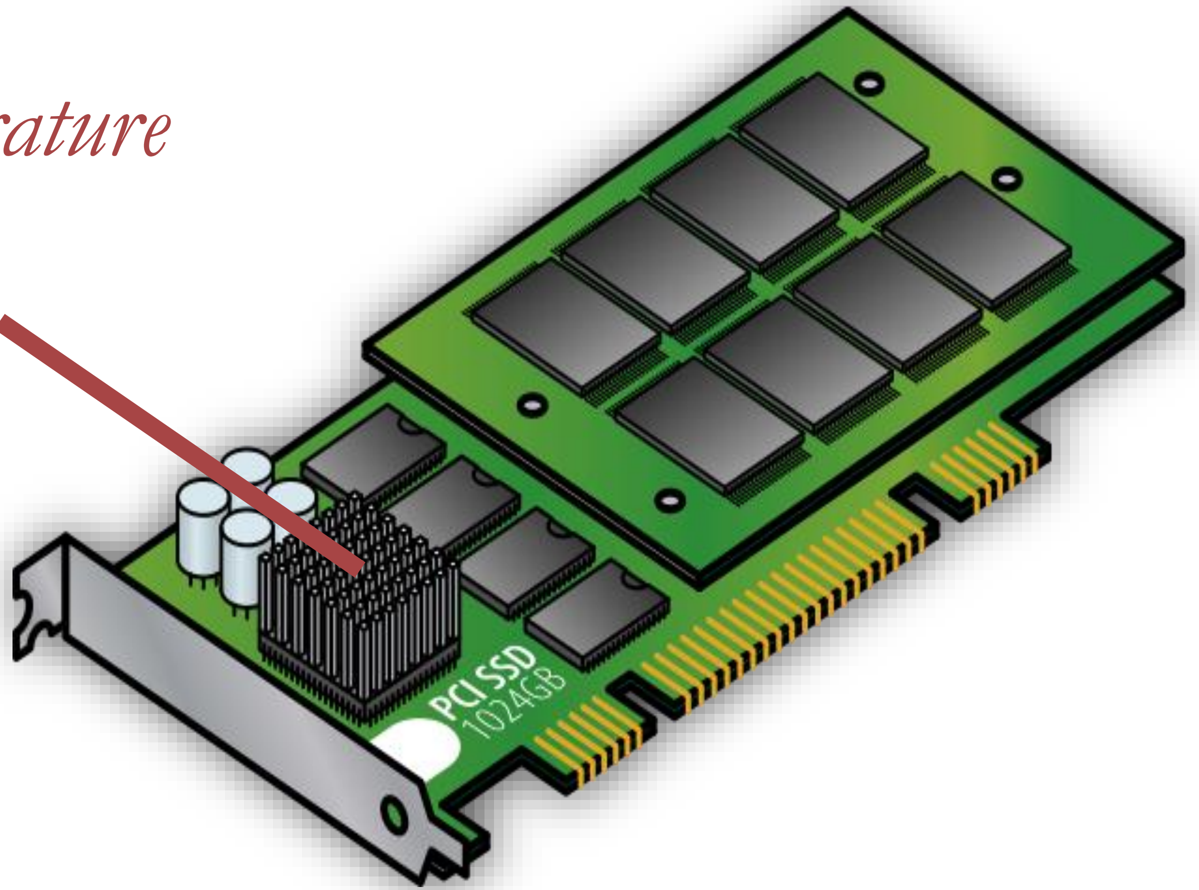
New reliability trends

Read disturbance

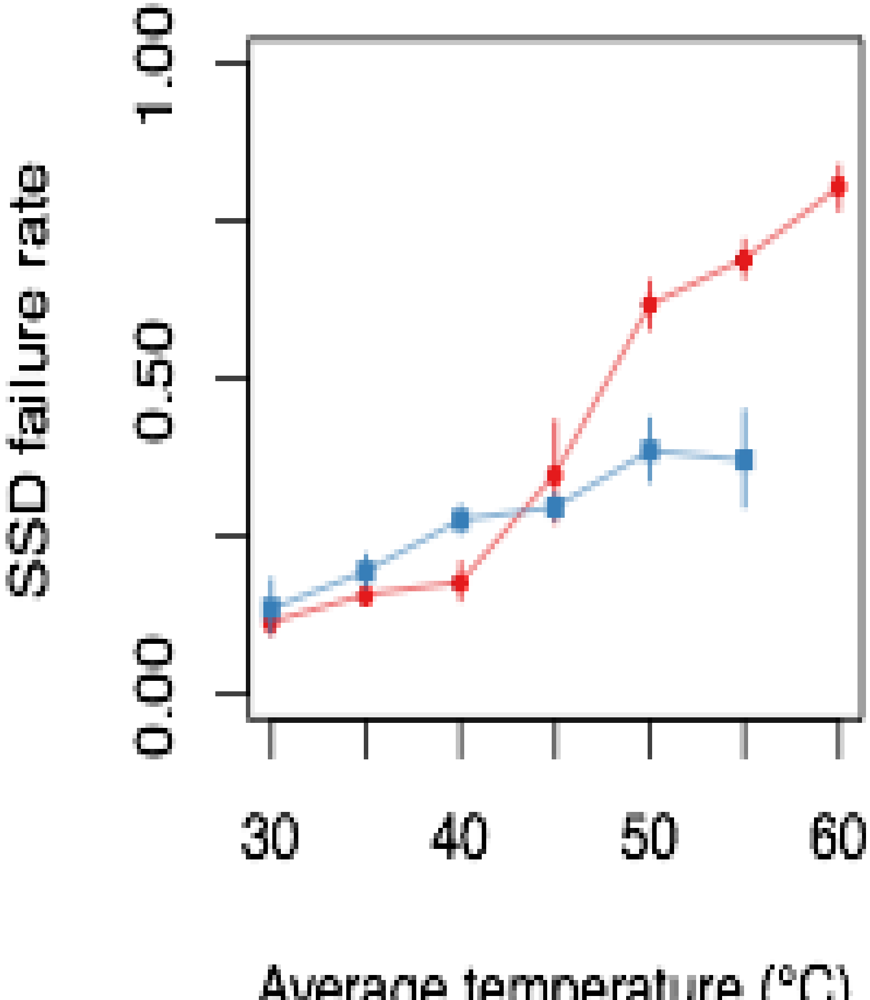
Temperature


PCI SSD
1024GB

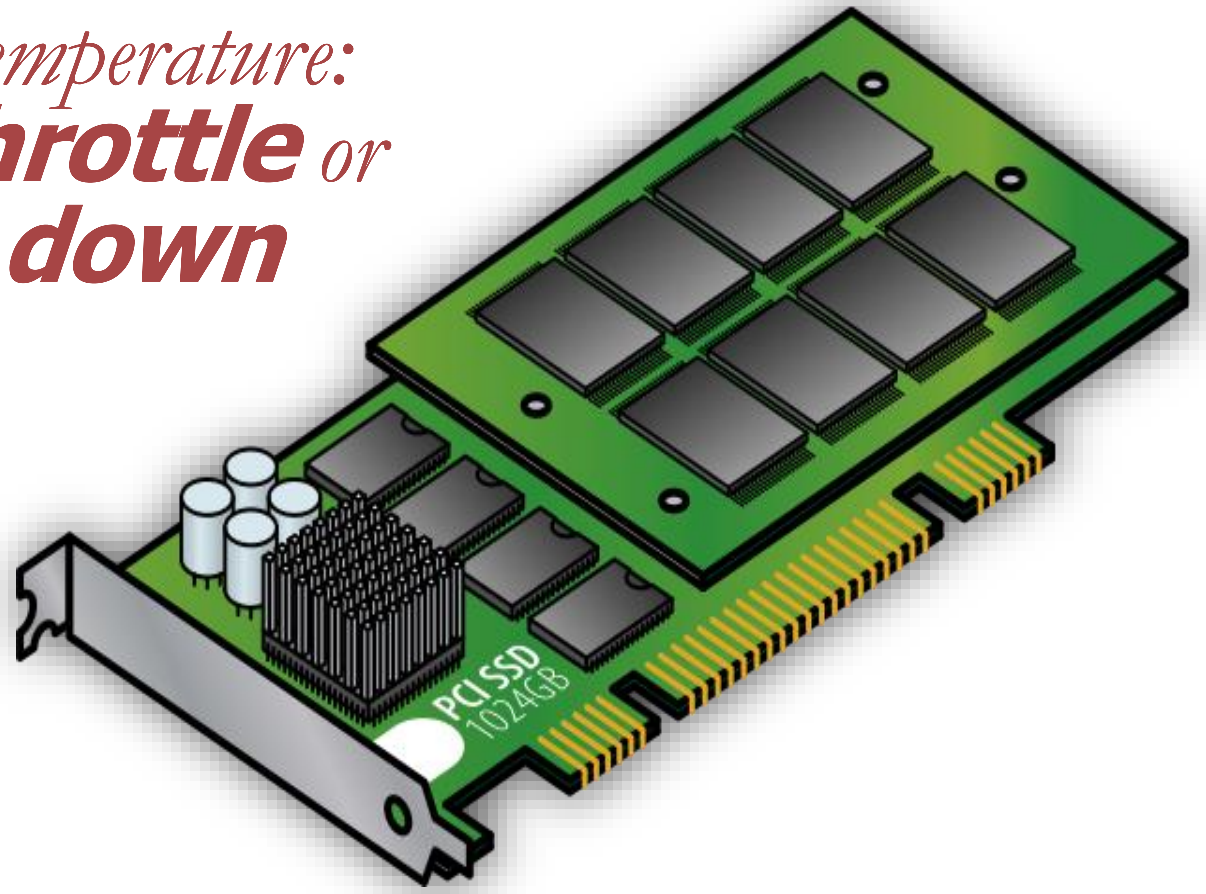
Temperature sensor



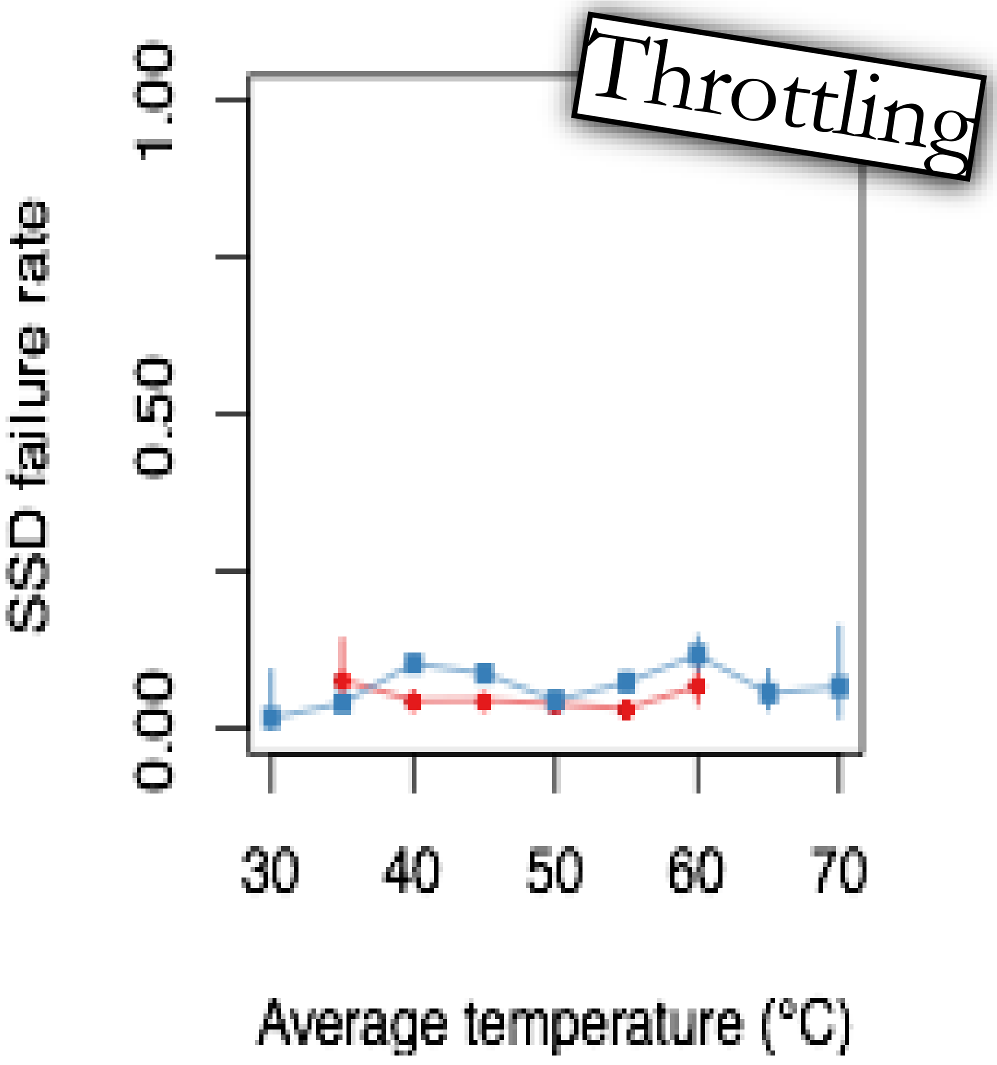
720GB, 1 SSD   720GB, 2 SSDs

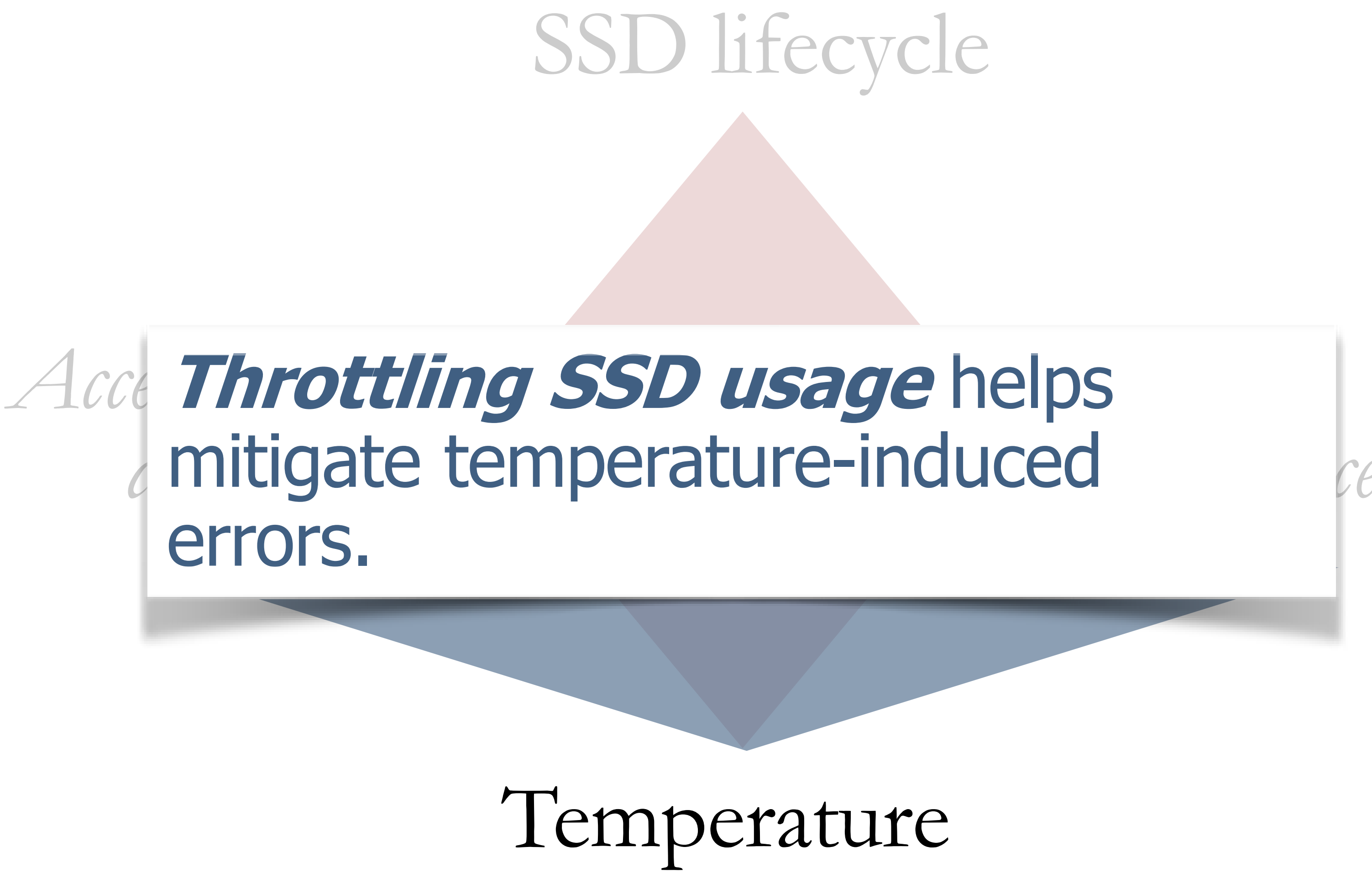


High temperature:
may **throttle** *or*
shut down



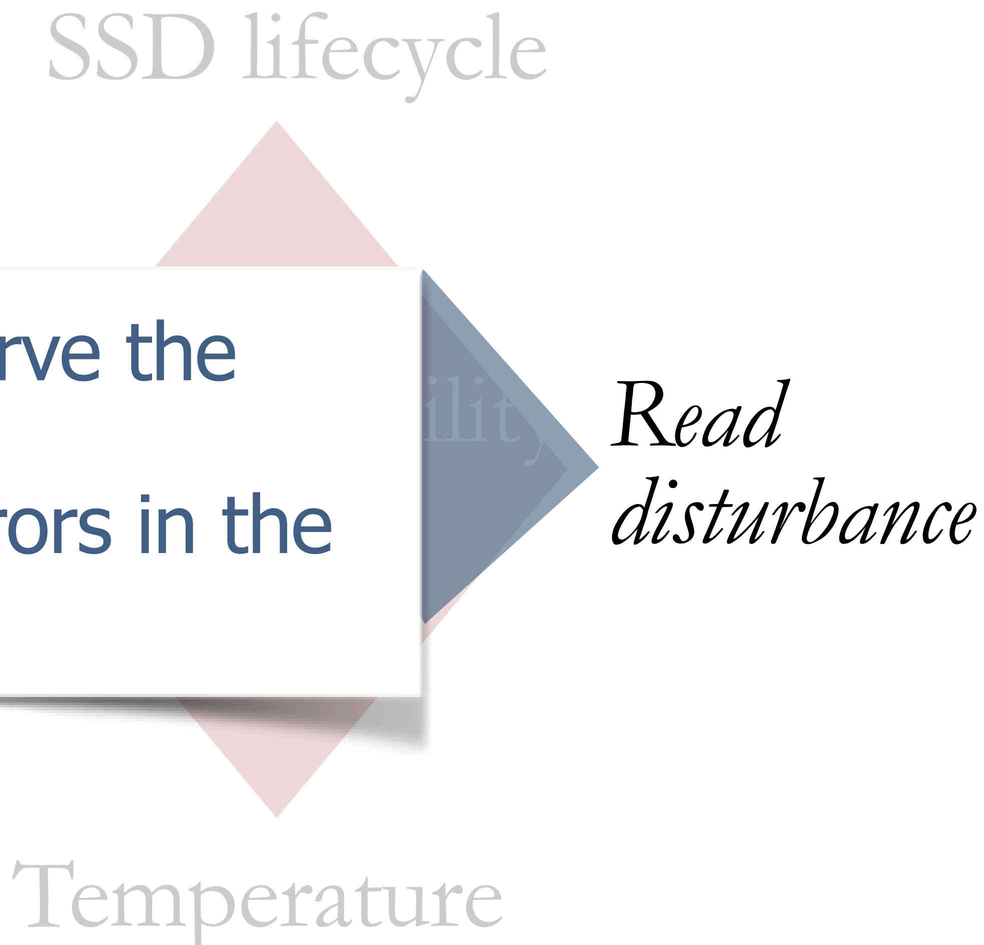
1.2TB, 1 SSD   3.2TB, 1 SSD





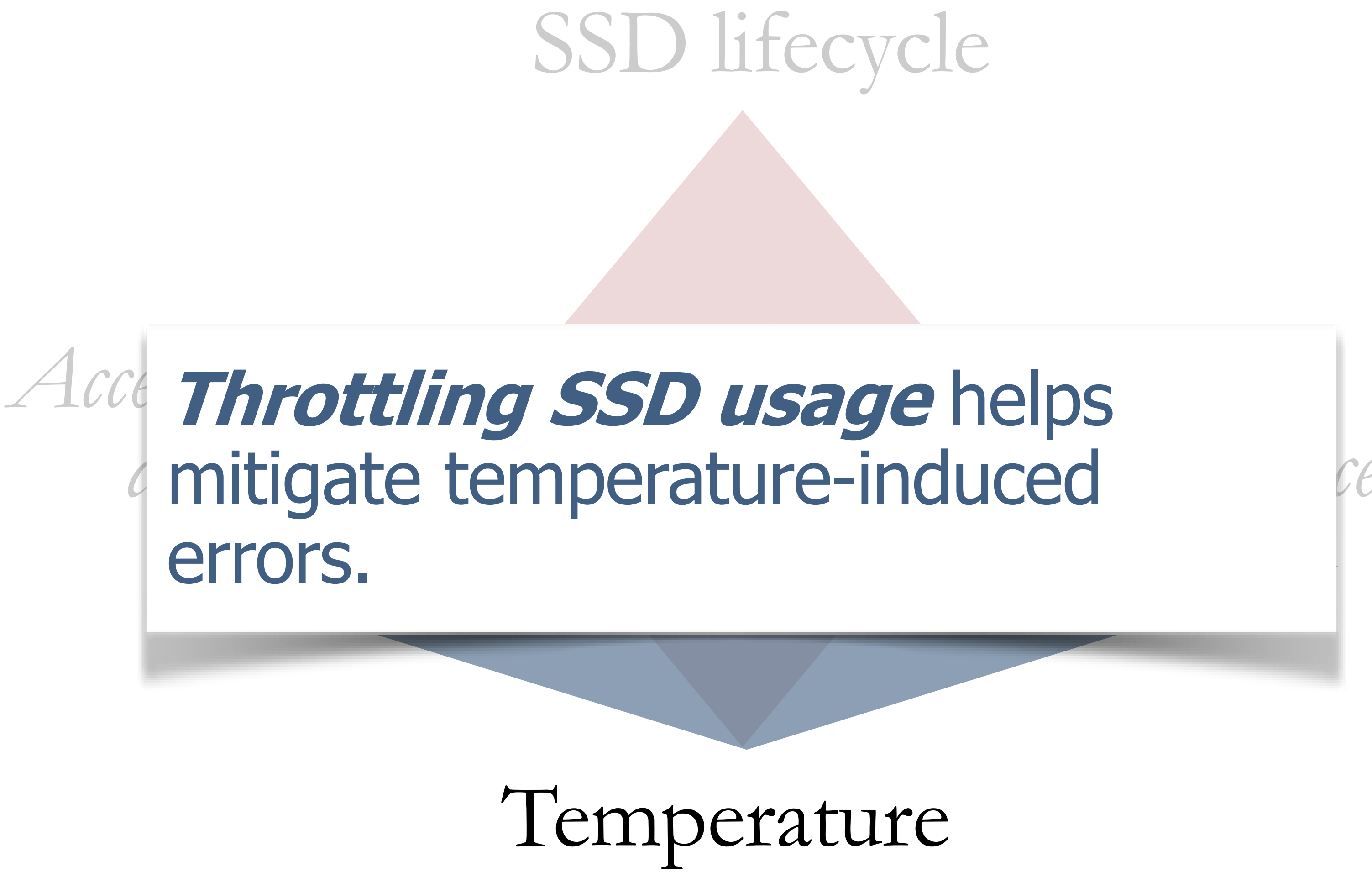
Throttling SSD usage helps mitigate temperature-induced errors.

Summary



We ***do not*** observe the effects of ***read disturbance*** errors in the field.

Summary



Throttling SSD usage helps mitigate temperature-induced errors.

Summary

SSD lifecycle

Access pattern dependence

We quantify the effects of the ***page cache*** and ***write amplification*** in the field.

Temperature

Large-Scale SSD Error Analysis [SIGMETRICS'15]

- First large-scale field study of flash memory errors

- Justin Meza, Qiang Wu, Sanjeev Kumar, and Onur Mutlu,
"A Large-Scale Study of Flash Memory Errors in the Field"
*Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (**SIGMETRICS**)*, Portland, OR, June 2015.
[Slides (pptx) (pdf)] [Coverage at ZDNet] [Coverage on The Register] [Coverage on TechSpot] [Coverage on The Tech Report]

A Large-Scale Study of Flash Memory Failures in the Field

Justin Meza
Carnegie Mellon University
meza@cmu.edu

Qiang Wu
Facebook, Inc.
qwu@fb.com

Sanjeev Kumar
Facebook, Inc.
skumar@fb.com

Onur Mutlu
Carnegie Mellon University
onur@cmu.edu

Other Works on NAND Flash Memory Modeling & Issues

Flash Memory Programming Vulnerabilities

- Yu Cai, Saugata Ghose, Yixin Luo, Ken Mai, Onur Mutlu, and Erich F. Haratsch,
  "Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques"
  *Proceedings of the 23rd International Symposium on High-Performance Computer Architecture (**HPCA**) Industrial Session*, Austin, TX, USA, February 2017.
  [Slides (pptx) (pdf)] [Lightning Session Slides (pptx) (pdf)]

Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques

Yu Cai[†] Saugata Ghose[†] Yixin Luo[‡†] Ken Mai[†] Onur Mutlu[§†] Erich F. Haratsch[‡]
[†]*Carnegie Mellon University* [‡]*Seagate Technology* [§]*ETH Zürich*

Accurate and Online Channel Modeling

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu, **"Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory"** *to appear in IEEE Journal on Selected Areas in Communications (**JSAC**)*, 2016.

Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory

Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, Onur Mutlu

Agenda

- Background, Motivation and Approach
- Experimental Characterization Methodology
- Error Analysis and Management
  - Main Characterization Results
  - Retention-Aware Error Management
  - Threshold Voltage and Program Interference Analysis
  - Read Reference Voltage Prediction
  - Neighbor-Assisted Error Correction
  - Read Disturb Error Handling
  - Retention Error Handling
  - Large Scale Field Analysis
  - 3D NAND Flash Memory Reliability
- Summary

3D NAND Flash Memory

3D NAND Flash Reliability I **[HPCA'18]**

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu, **"HeatWatch: Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature-Awareness"**
*Proceedings of the 24th International Symposium on High-Performance Computer Architecture (**HPCA**)*, Vienna, Austria, February 2018.
[Lightning Talk Video]
[Slides (pptx) (pdf)] [Lightning Session Slides (pptx) (pdf)]

HeatWatch: Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature Awareness

Yixin Luo[†] Saugata Ghose[†] Yu Cai[‡] Erich F. Haratsch[‡] Onur Mutlu[§†]
[†]*Carnegie Mellon University* [‡]*Seagate Technology* [§]*ETH Zürich*

3D NAND Flash Reliability II [SIGMETRICS'18]

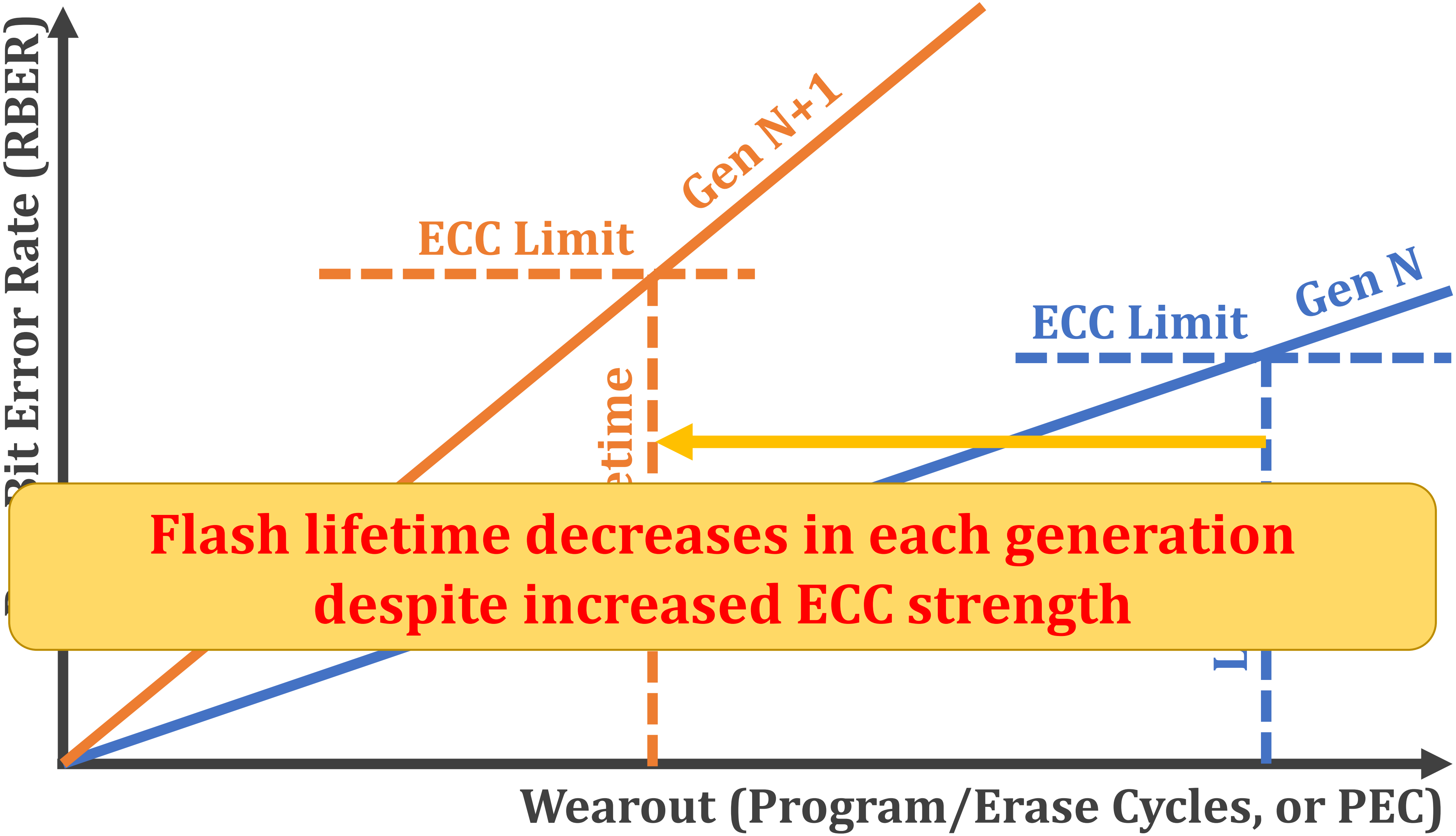
- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu,
"Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation"
*Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (**SIGMETRICS**)*, Irvine, CA, USA, June 2018.
[Abstract]

Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation

Yixin Luo† Saugata Ghose† Yu Cai† Erich F. Haratsch‡ Onur Mutlu§†

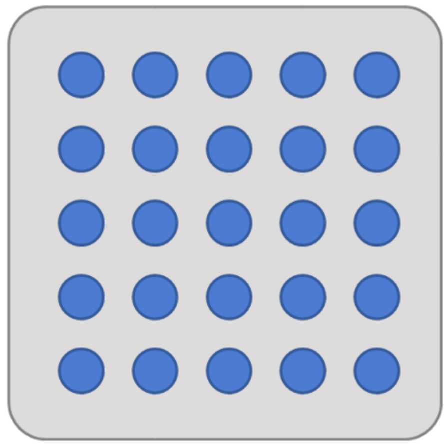
†Carnegie Mellon University ‡Seagate Technology §ETH Zürich

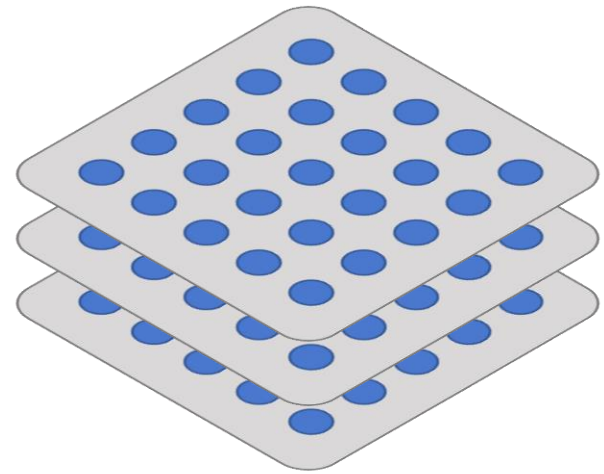
NAND Flash Memory Lifetime Problem



Flash lifetime decreases in each generation despite increased ECC strength

Planar vs. 3D NAND Flash Memory

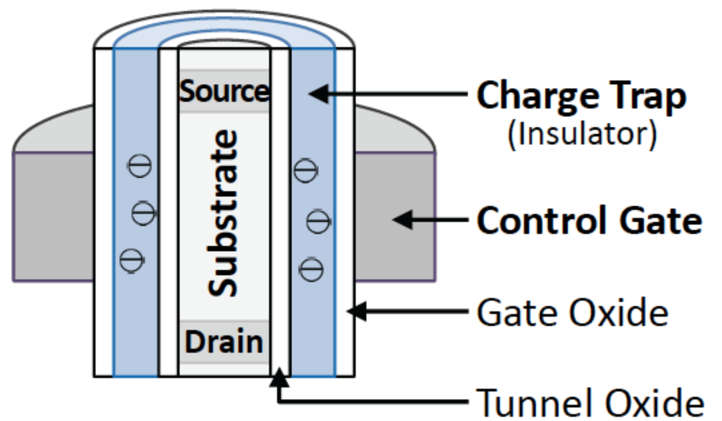




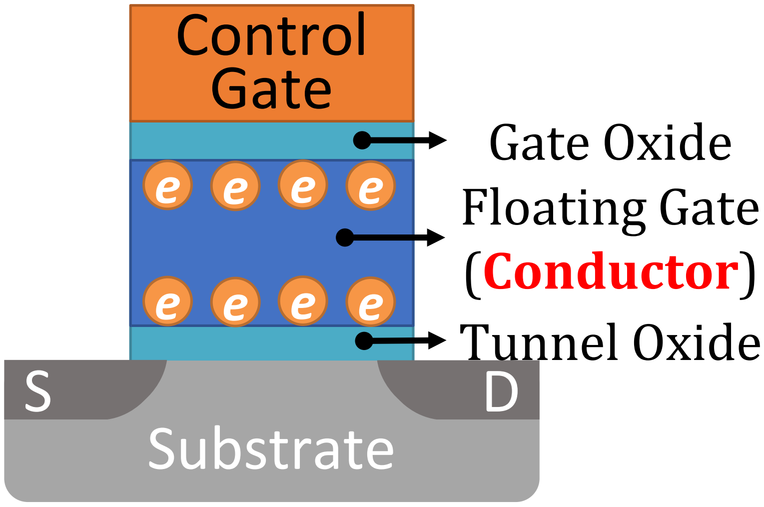
	Planar NAND Flash Memory	3D NAND Flash Memory
Scaling	Reduce flash cell size, Reduce distance b/w cells	Increase # of layers
Reliability	Scaling hurts reliability	**Not well studied!**

Charge Trap Based 3D Flash Cell

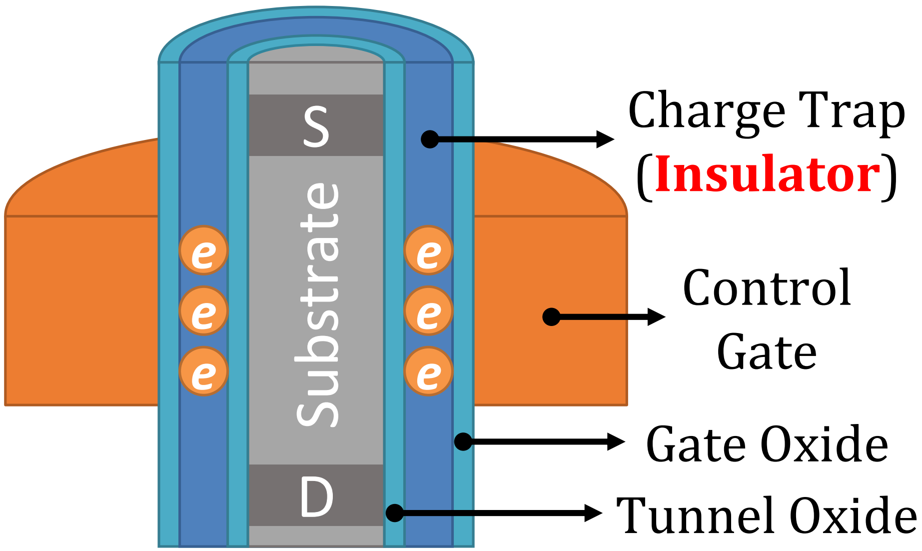
- Cross-section of a charge trap transistor



2D vs. 3D Flash Cell Design



2D Floating-Gate Cell



3D Charge-Trap Cell

3D NAND Flash Memory Organization

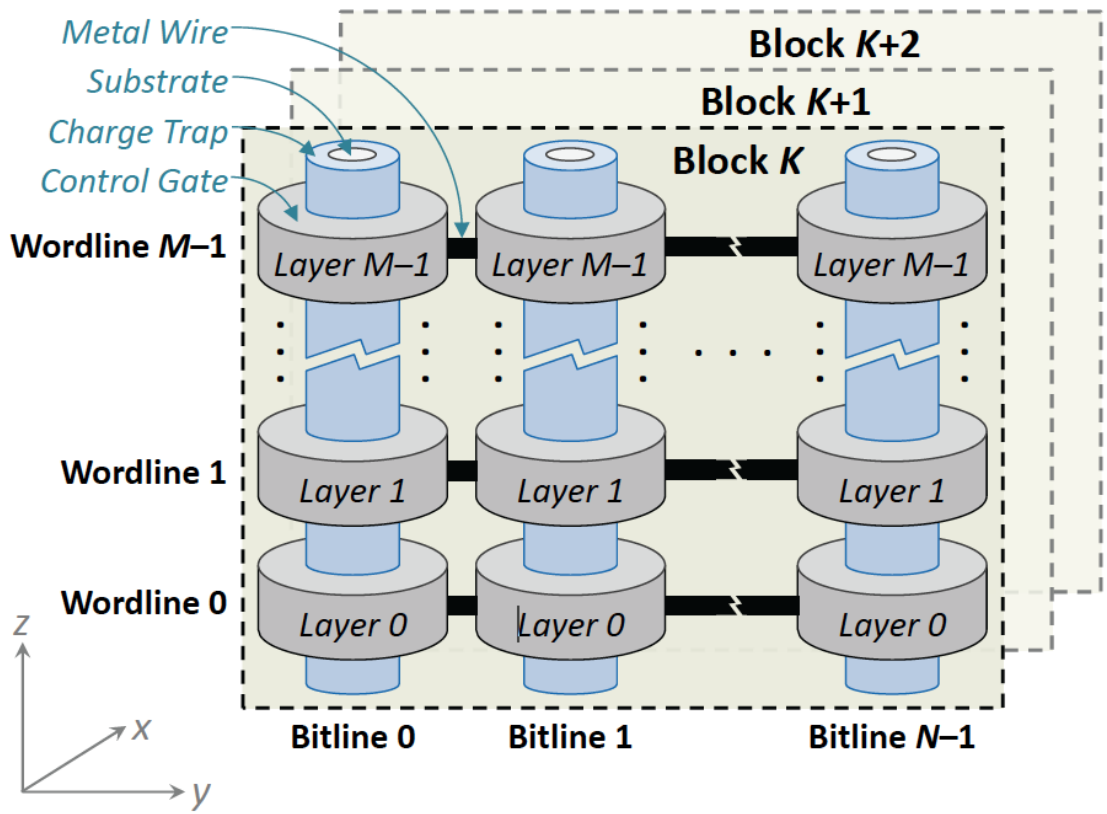


Fig. 43. Organization of flash cells in an *M*-layer 3D charge trap NAND flash memory chip, where each block consists of *M* wordlines and *N* bitlines.

More Background and State-of-the-Art

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu, **"Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery"**
  Invited Book Chapter in Inside Solid State Drives, 2018.
  [Preliminary arxiv.org version]

Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery

YU CAI, SAUGATA GHOSE
Carnegie Mellon University

ERICH F. HARATSCH
Seagate Technology

YIXIN LUO
Carnegie Mellon University

ONUR MUTLU
ETH Zürich and Carnegie Mellon University

3D vs. Planar NAND Errors: Comparison

Table 4. Changes in behavior of different types of errors in 3D NAND flash memory, compared to planar (i.e., two-dimensional) NAND flash memory. See Section 6.2 for a detailed discussion.

Error Type	Change in 3D vs. Planar
P/E Cycling (Section 3.1)	3D is *less susceptible*, due to current use of charge trap transistors for flash cells
Program (Section 3.2)	3D is *less susceptible for now*, due to use of one-shot programming (see Section 2.4)
Cell-to-Cell Interference (Section 3.3)	3D is *less susceptible for now*, due to larger manufacturing process technology
Data Retention (Section 3.4)	3D is *more susceptible*, due to early retention loss
Read Disturb (Section 3.5)	3D is *less susceptible for now*, due to larger manufacturing process technology

Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation

Yixin Luo Saugata Ghose Yu Cai Erich F. Haratsch Onur Mutlu

Carnegie Mellon SK hynix ETH Zürich
SAFARI SEAGATE

Executive Summary

- Problem: 3D NAND error characteristics are **not well studied**
- Goal: ***Understand*** & ***mitigate*** 3D NAND errors to improve lifetime
- **Contribution 1: Characterize** real 3D NAND flash chips
  - ***Process variation:*** ***21×*** *error rate difference across layers*
  - ***Early retention loss:*** *Error rate increases by* ***10×*** *after 3 hours*
  - ***Retention interference:*** ***Not observed before*** *in planar NAND*
- **Contribution 2: Model** RBER and threshold voltage
  - *RBER (raw bit error rate) variation model*
  - *Retention loss model*
- **Contribution 3: Mitigate** 3D NAND flash errors
  - ***LaVAR:*** *Layer Variation Aware Reading*
  - ***LI-RAID:*** *Layer-Interleaved RAID*
  - ***ReMAR:*** *Retention Model Aware Reading*
  - *Improve flash lifetime by* ***1.85×*** *or reduce ECC overhead by* ***78.9%***

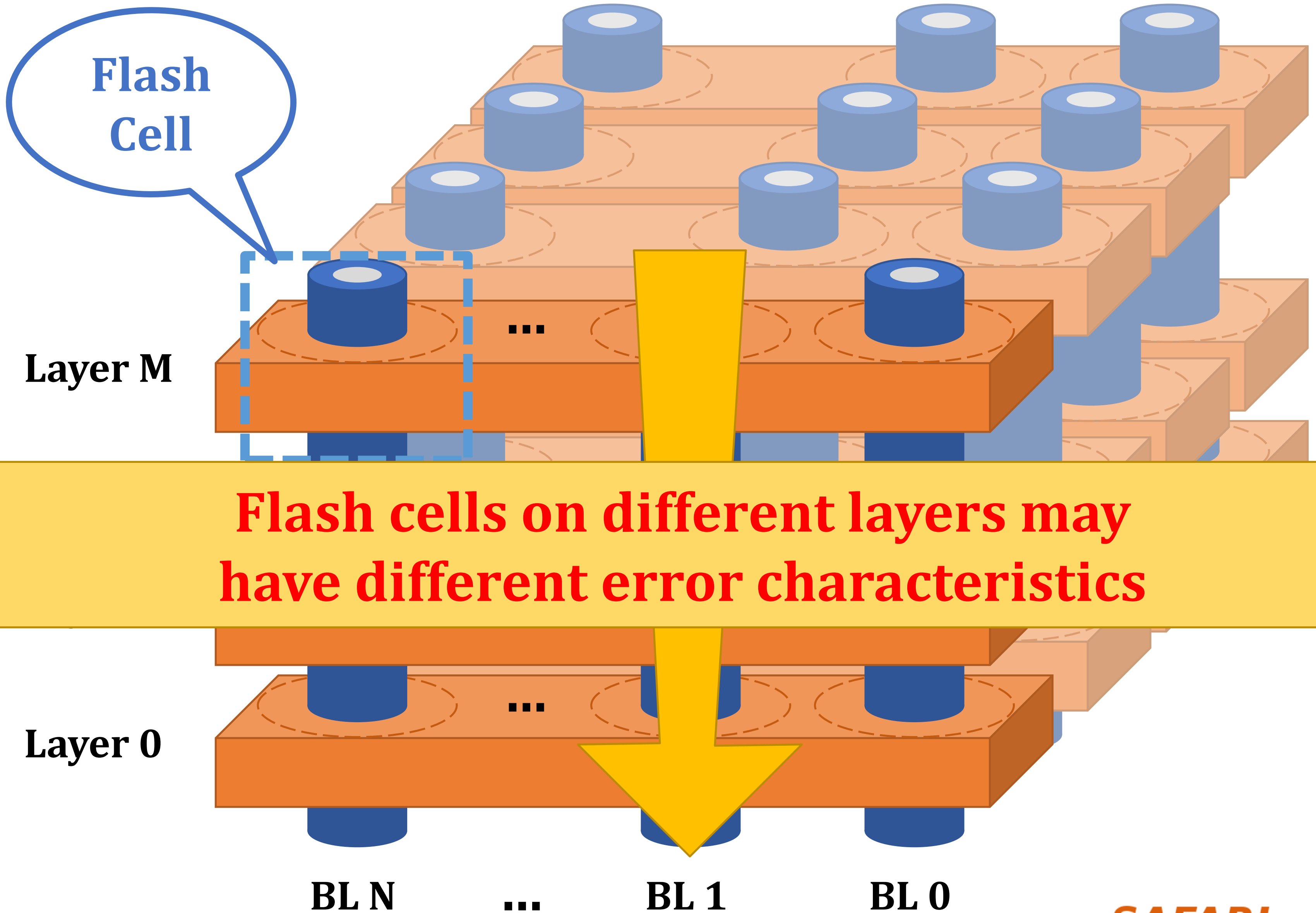
Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
- Contribution 2: Model RBER and threshold voltage
- Contribution 3: Mitigate 3D NAND flash errors
- Conclusion

Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
  - Process variation
  - Early retention loss
  - Retention interference
- Contribution 2: Model RBER and threshold voltage
- Contribution 3: Mitigate 3D NAND flash errors
- Conclusion

Process Variation Across Layers

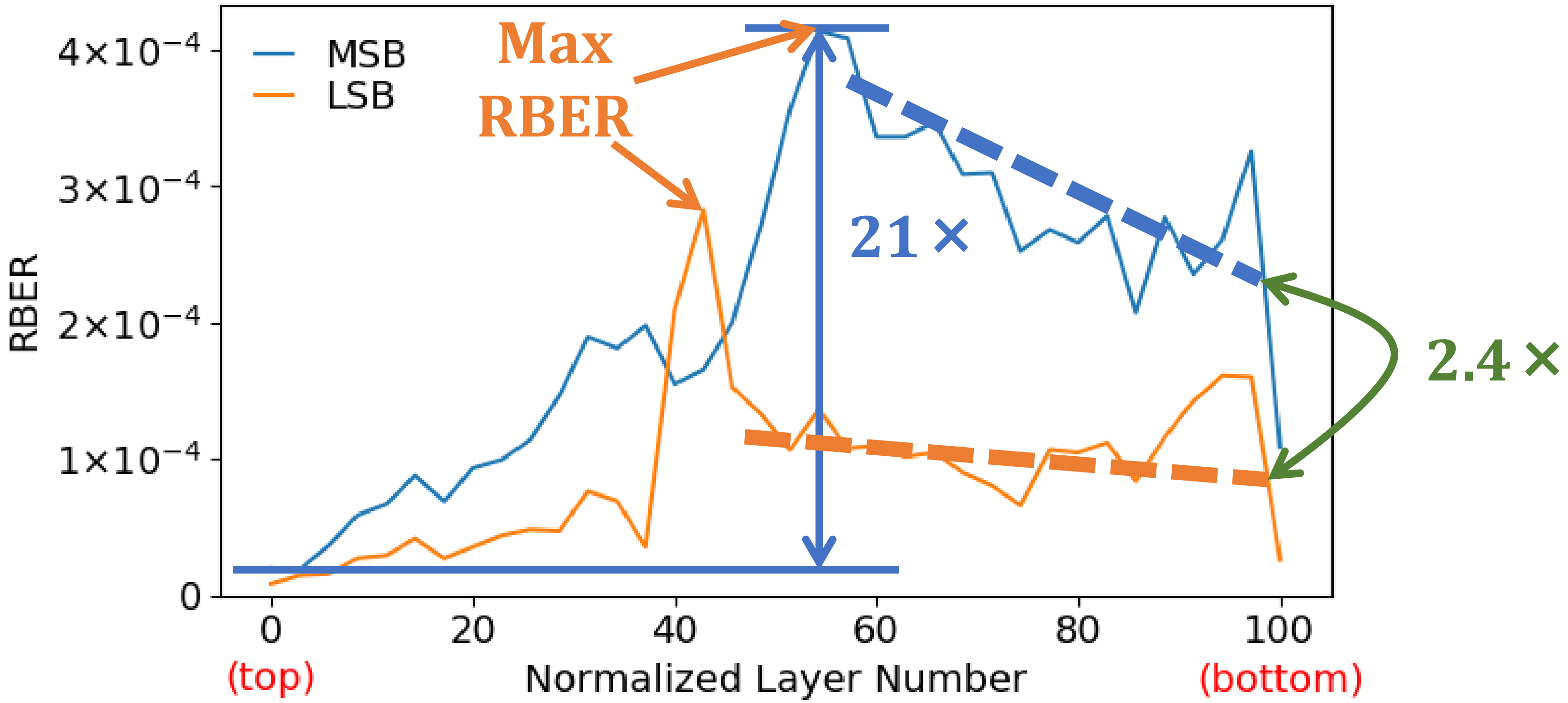


Flash cells on different layers may have different error characteristics

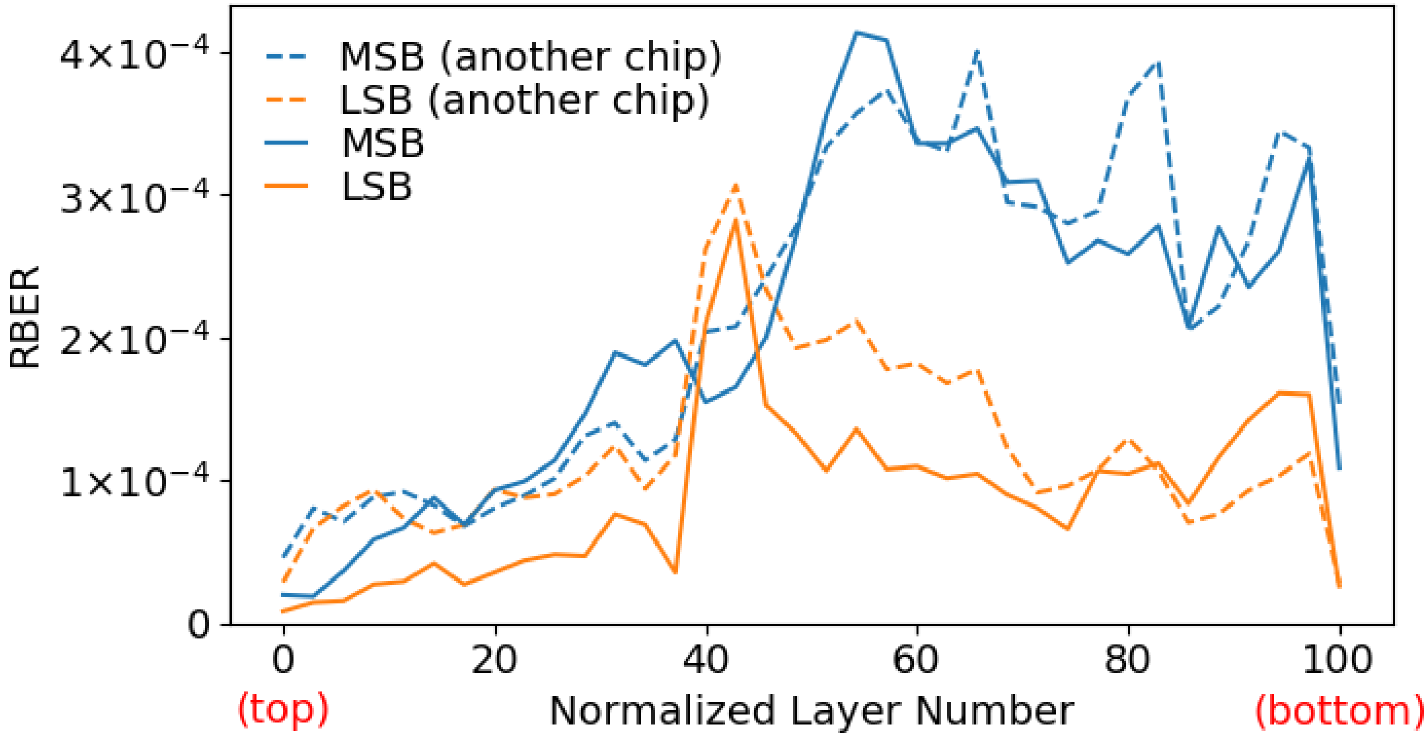
Characterization Methodology

- Modified firmware version in the flash controller
  - Controls the read reference voltage of the flash chip
  - Bypasses ECC to get raw data (with raw bit errors)
- Analysis and post-processing of the data on the server

Layer-to-Layer Process Variation



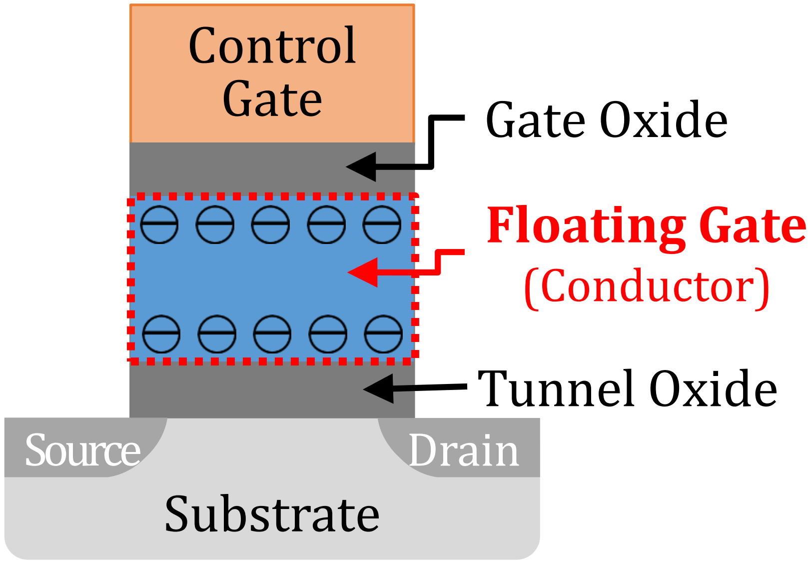
Layer-to-Layer Process Variation



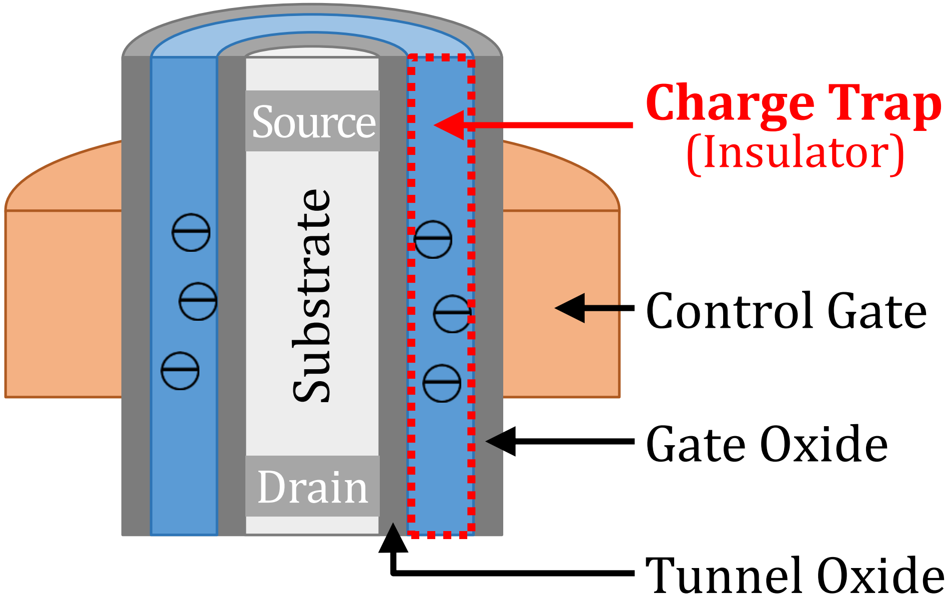
Large RBER variation across layers and LSB-MSB pages

Retention Loss Phenomenon

Planar NAND Cell

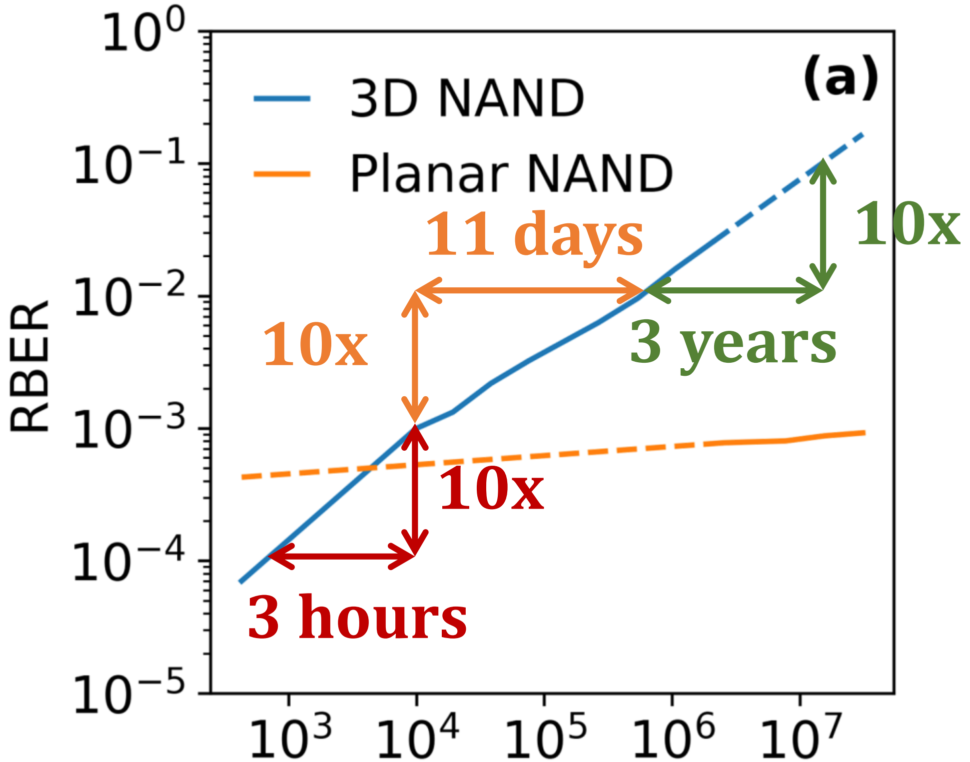


3D NAND Cell



Most dominant type of error in planar NAND.
Is this true for 3D NAND as well?

Early Retention Loss



Retention errors increase quickly immediately after programming

Characterization Summary

- **Layer-to-layer process variation**
  - Large RBER variation across layers and LSB-MSB pages
  - → Need new mechanisms to tolerate RBER variation!
- **Early retention loss**
  - RBER increases quickly after programming
  - → Need new mechanisms to tolerate retention errors!
- **Retention interference**
  - Amount of retention loss correlated with neighbor cells' states
  - → Need new mechanisms to tolerate retention interference!

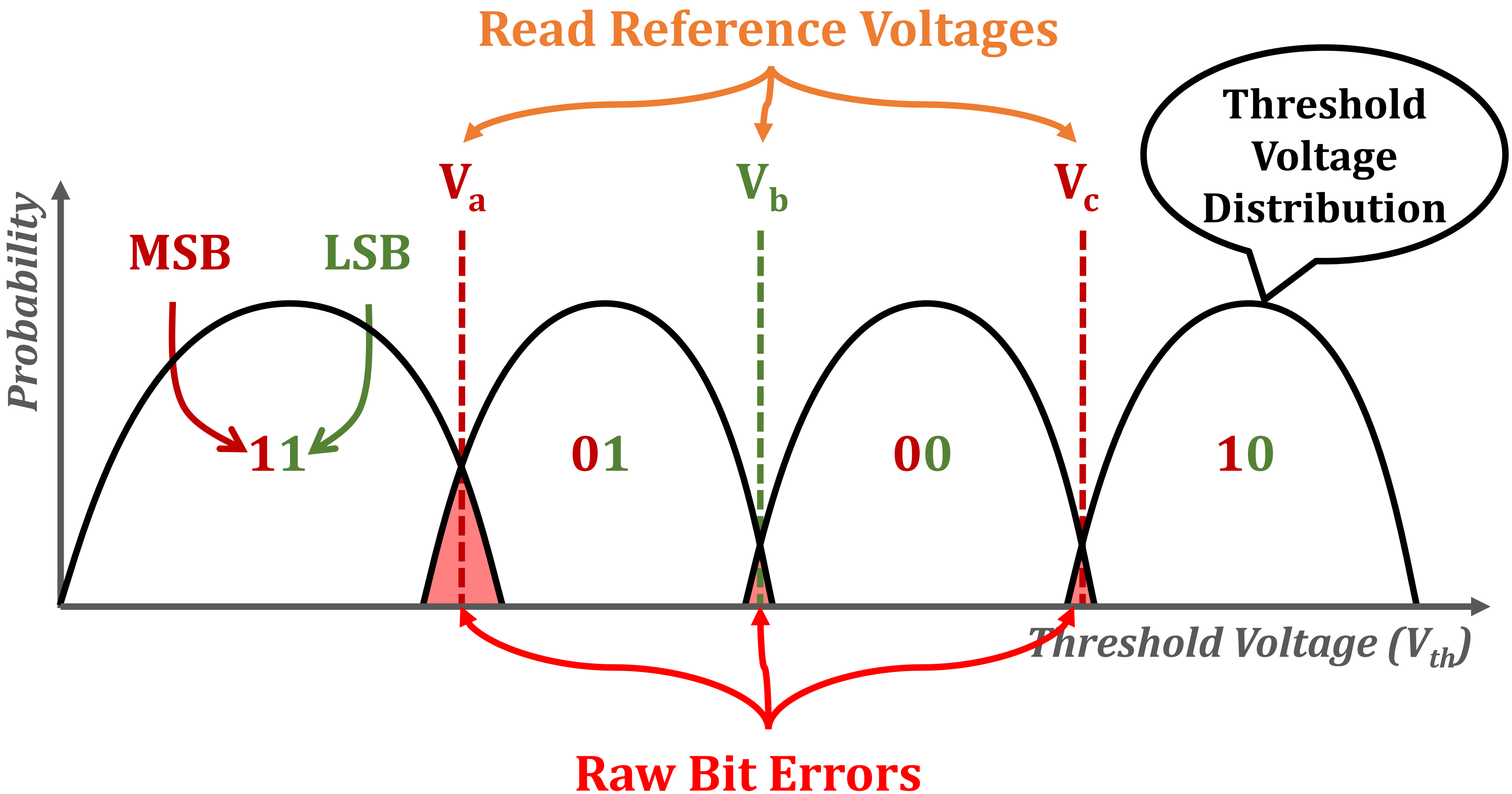
- **More *threshold voltage* and *RBER* results in the paper:** 3D NAND P/E cycling, program interference, read disturb, read variation, bitline-to-bitline process variation

- **Our approach** based on insights developed via our experimental characterization: Develop **error models**, and build online **error mitigation mechanisms** using the models

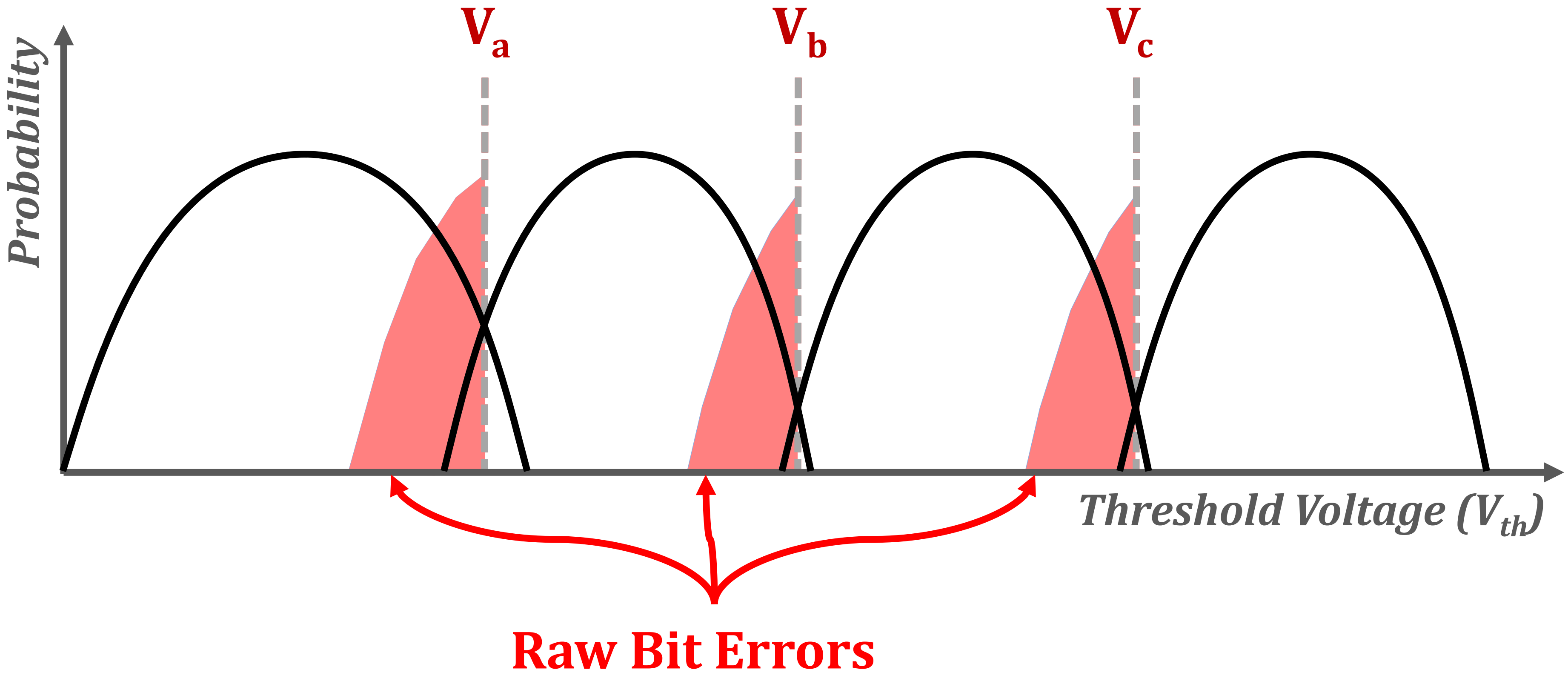
Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
- Contribution 2: Model RBER and threshold voltage
  - Retention loss model
  - RBER variation model
- Contribution 3: Mitigate 3D NAND flash errors
- Conclusion

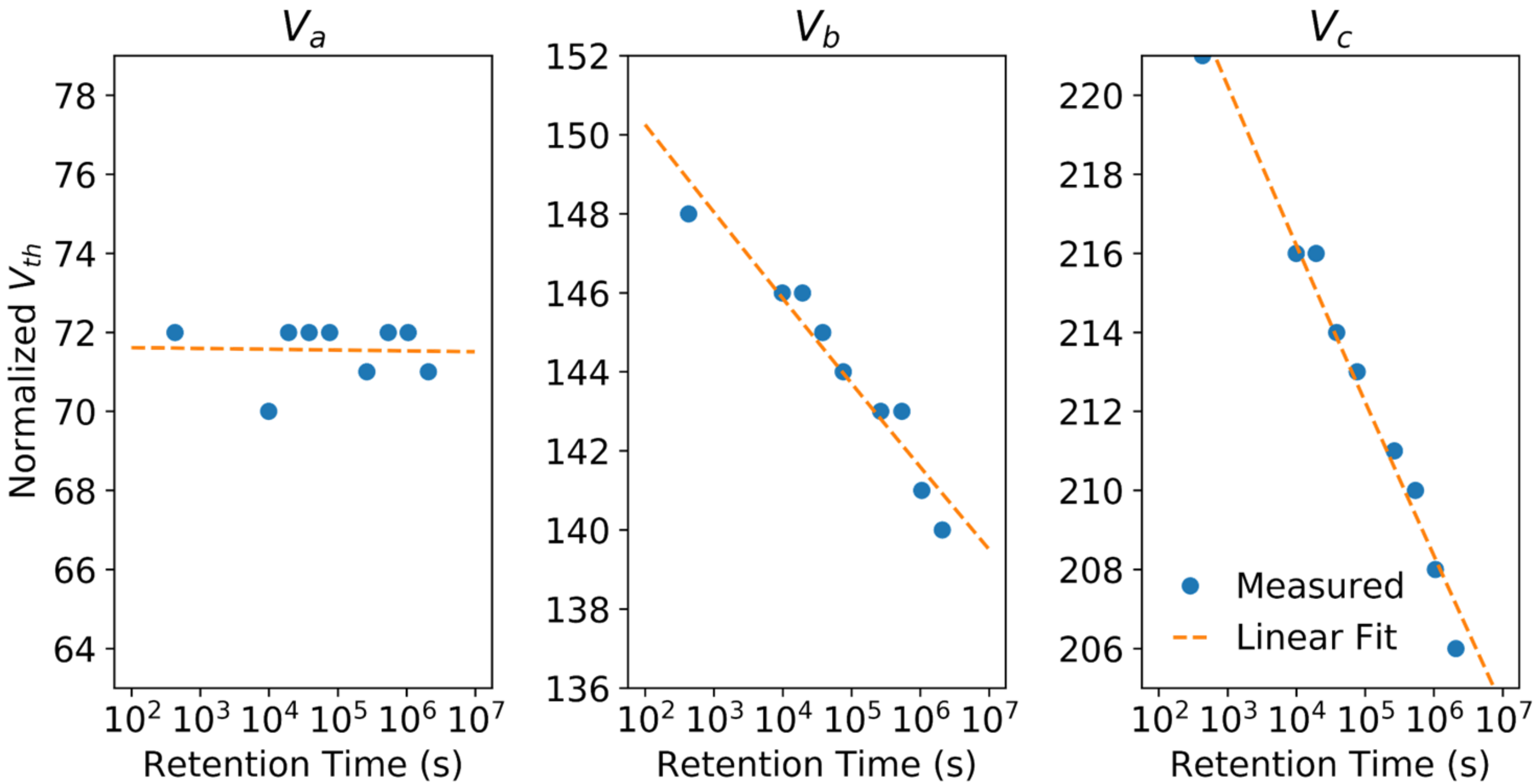
What Do We Model?



Optimal Read Reference Voltage



Retention Loss Model



Early retention loss can be modeled as a simple linear function of log(retention time)

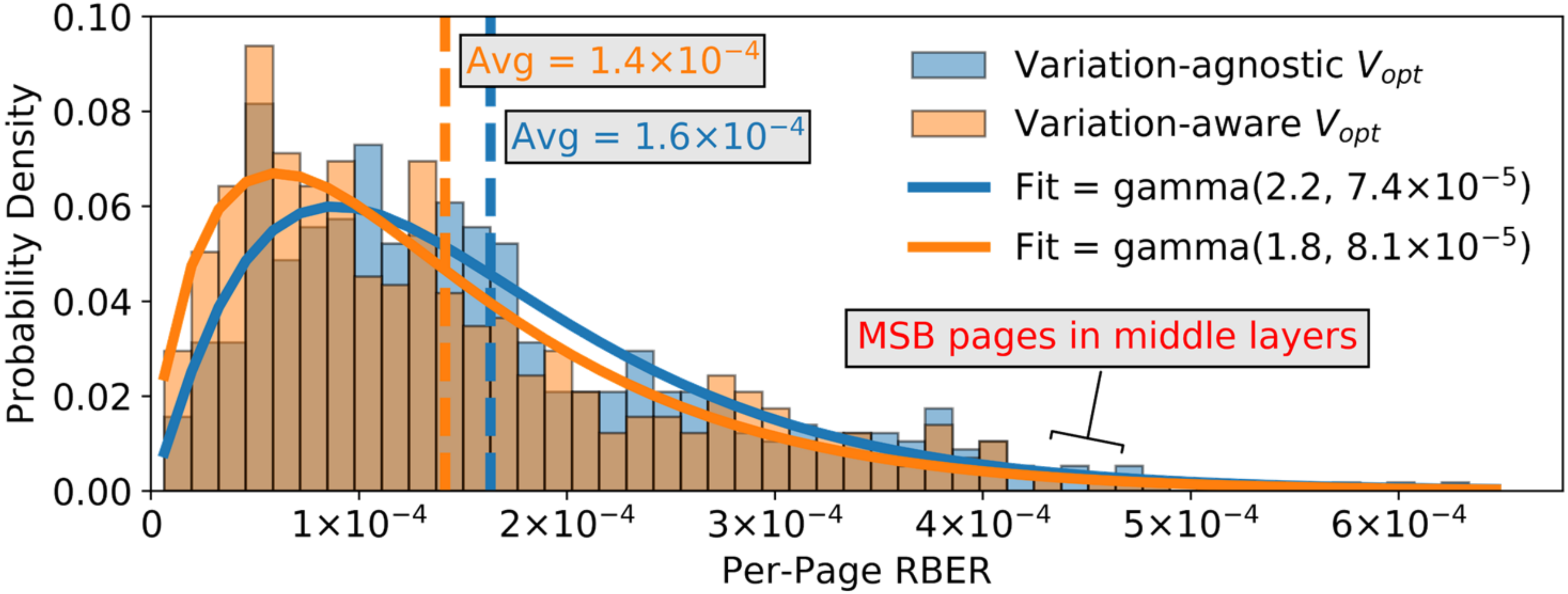
Retention Loss Model

- Goal: Develop a simple linear model that can be used online

- **Models**
  - Optimal read reference voltage (V_b and V_c)
  - Raw bit error rate ($log(RBER)$)
  - Mean and standard deviation of threshold voltage distribution (μ and σ)
- **As a function of**
  - Retention time ($log(t)$)
  - P/E cycle count (PEC)
- e.g., $V_{opt} = (\alpha \times PEC + \beta) \times log(t) + \gamma \times PEC + \delta$

- Model error <1 step for V_b and V_c
- Adjusted R^2 > 89%

RBER Variation Model



Variation-agnostic V_{opt}

- Same V_{ref} for all layers optimized for the entire block

RBER distribution follows gamma distribution

KL-divergence error = 0.09

SAFARI

Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
- Contribution 2: Model RBER and threshold voltage
- Contribution 3: Mitigate 3D NAND flash errors
  - LaVAR: Layer Variation Aware Reading
  - LI-RAID: Layer-Interleaved RAID
  - ReMAR: Retention Model Aware Reading
- Conclusion

LaVAR: Layer Variation Aware Reading

- **Layer-to-layer process variation**
  - Error characteristics are different in each layer
- **Goal**: Adjust read reference voltage **for each layer**

- **Key Idea**: Learn a ***voltage offset (Offset)*** for each layer
  - $V_{opt}^{Layer\ aware} = V_{opt}^{Layer\ agnostic} + Offset$

- **Mechanism**
  - **Offset**: Learned once for each chip & stored in a table
    - *Uses (**2 × Layers**) Bytes memory per chip*
  - $V_{opt}^{Layer\ agnostic}$: Predicted by any existing V_{opt} model
    - *E.g., ReMAR [Luo+Sigmetrics'18], HeatWatch [Luo+HPCA'18], OFCM [Luo+JSAC'16], ARVT [Papandreou+GLSVLSI'14]*

- Reduces RBER on average by **43%** (based on our characterization data)

LI-RAID: Layer-Interleaved RAID

- **Layer-to-layer process variation**
  - Worst-case RBER much higher than average RBER
- **Goal:** Significantly reduce worst-case RBER

- **Key Idea**
  - Group flash pages on ***less reliable layers*** with pages on ***more reliable layers***
  - Group ***MSB pages*** with ***LSB pages***

- **Mechanism**
  - Reorganize RAID layout to eliminate worst-case RBER
  - **<0.8%** storage overhead

Conventional RAID

Wordline #	*Layer #*	*Page*	**Chip 0**	**Chip 1**	**Chip 2**	**Chip 3**
0	*0*	*MSB*	Group 0	Group 0	Group 0	Group 0
0	*0*	*LSB*	Group 1	Group 1	Group 1	Group 1
1	*1*	*MSB*	Group 2	Group 2	Group 2	Group 2
1	*1*	*LSB*	Group 3	Group 3	Group 3	Group 3
2	*2*	*MSB*	Group 4	Group 4	Group 4	Group 4
2	*2*	*LSB*	Group 5	Group 5	Group 5	Group 5
3	*3*	*MSB*	Group 6	Group 6	Group 6	Group 6
3	*3*	*LSB*	Group 7	Group 7	Group 7	Group 7

Worst-case RBER in any layer limits the lifetime of conventional RAID

LI-RAID: Layer-Interleaved RAID

Wordline #	*Layer #*	*Page*	Chip 0	Chip 1	Chip 2	Chip 3
0	*0*	*MSB*	Group 0	Blank	Group 4	Group 3
0	*0*	*LSB*	Group 1	Blank	Group 5	Group 2
1	*1*	*MSB*	Group 2	Group 1	Blank	Group 5
1	*1*	*LSB*	Group 3	Group 0	Blank	Group 4
2	*2*	*MSB*	Group 4	Group 3	Group 0	Blank
2	*2*	*LSB*	Group 5	Group 2	Group 1	Blank
3	*3*	*MSB*	Blank	Group 5	Group 2	Group 1
3	*3*	*LSB*	Blank	Group 4	Group 3	Group 0

Any page with worst-case RBER can be corrected by other reliable pages in the RAID group

LI-RAID: Layer-Interleaved RAID

- **Layer-to-layer process variation**
  - Worst-case RBER much higher than average RBER
- **Goal:** Significantly reduce worst-case RBER

- **Key Idea**
  - Group flash pages on ***less reliable layers*** with pages on ***more reliable layers***
  - Group ***MSB pages*** with ***LSB pages***

- **Mechanism**
  - Reorganize RAID layout to eliminate worst-case RBER
  - **<0.8%** storage overhead

- Reduces worst-case RBER by **66.9%** (based on our characterization data)

ReMAR: Retention Model Aware Reading

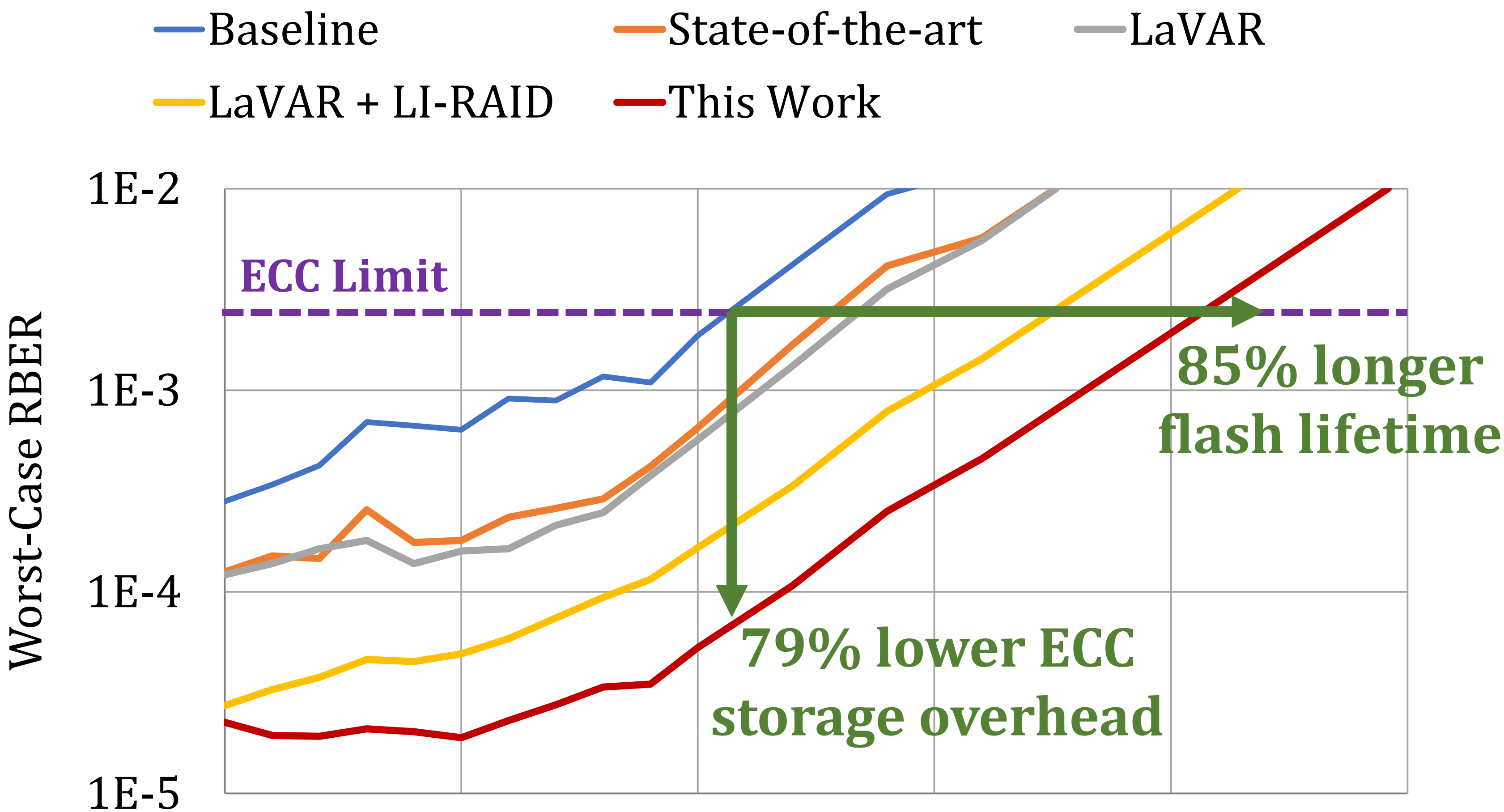
- **Early retention loss**
  - Threshold voltage shifts quickly after programming
- **Goal: Adjust read reference voltages based on retention loss**

- **Key Idea:** Learn and use a retention loss model online

- **Mechanism**
  - Periodically characterize and learn retention loss model online
  - Retention time = Read timestamp - Write timestamp
    - *Uses **800 KB** memory to store program time of each block*
  - Predict retention-aware V_{opt} using the model

- Reduces RBER on average by **51.9%**
  (based on our characterization data)

Impact on System Reliability



LaVAR, LI-RAID, and ReMAR improve flash lifetime or reduce ECC overhead significantly

Error Mitigation Techniques Summary

- **LaVAR: Layer Variation Aware Reading**
  - Learn a V_{opt} offset for each layer and apply ***layer-aware V_{opt}***
- **LI-RAID: Layer-Interleaved RAID**
  - Group flash pages on ***less reliable layers*** with pages on ***more reliable layers***
  - Group ***MSB pages*** with ***LSB pages***
- **ReMAR: Retention Model Aware Reading**
  - Learn retention loss model and apply ***retention-aware V_{opt}***
- **Benefits:**
- Improve flash lifetime by **1.85×** or reduce ECC overhead by **78.9%**

- **ReNAC (in paper):** Reread a failed page using V_{opt} based on the *retention interference* induced by neighbor cell

Agenda

- Background & Introduction
- Contribution 1: Characterize real 3D NAND flash chips
- Contribution 2: Model RBER and threshold voltage
- Contribution 3: Mitigate 3D NAND flash errors
- Conclusion

Conclusion

- Problem: 3D NAND error characteristics are **not well studied**
- Goal: ***Understand*** & ***mitigate*** 3D NAND errors to improve lifetime
- **Contribution 1: Characterize** real 3D NAND flash chips
  - ***Process variation:*** ***21×*** *error rate difference across layers*
  - ***Early retention loss:*** *Error rate increases by* ***10×*** *after 3 hours*
  - ***Retention interference:*** ***Not observed before*** *in planar NAND*
- **Contribution 2: Model** RBER and threshold voltage
  - *RBER (raw bit error rate) variation model*
  - *Retention loss model*
- **Contribution 3: Mitigate** 3D NAND flash errors
  - ***LaVAR:*** *Layer Variation Aware Reading*
  - ***LI-RAID:*** *Layer-Interleaved RAID*
  - ***ReMAR:*** *Retention Model Aware Reading*
  - *Improve flash lifetime by* ***1.85×*** *or reduce ECC overhead by* ***78.9%***

Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation

Yixin Luo Saugata Ghose Yu Cai Erich F. Haratsch Onur Mutlu

3D NAND Flash Reliability II [SIGMETRICS'18]

- Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu, **"Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation"** *Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems* (***SIGMETRICS***), Irvine, CA, USA, June 2018.
[Abstract]

Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation

Yixin Luo[†] Saugata Ghose[†] Yu Cai[†] Erich F. Haratsch[‡] Onur Mutlu[§†]

[†]Carnegie Mellon University [‡]Seagate Technology [§]ETH Zürich

One More Idea

WARM

Improving NAND Flash Memory Lifetime with Write-hotness Aware Retention Management

***Yixin Luo**, Yu Cai, Saugata Ghose, Jongmoo Choi*, Onur Mutlu*

*Carnegie Mellon University, *Dankook University*

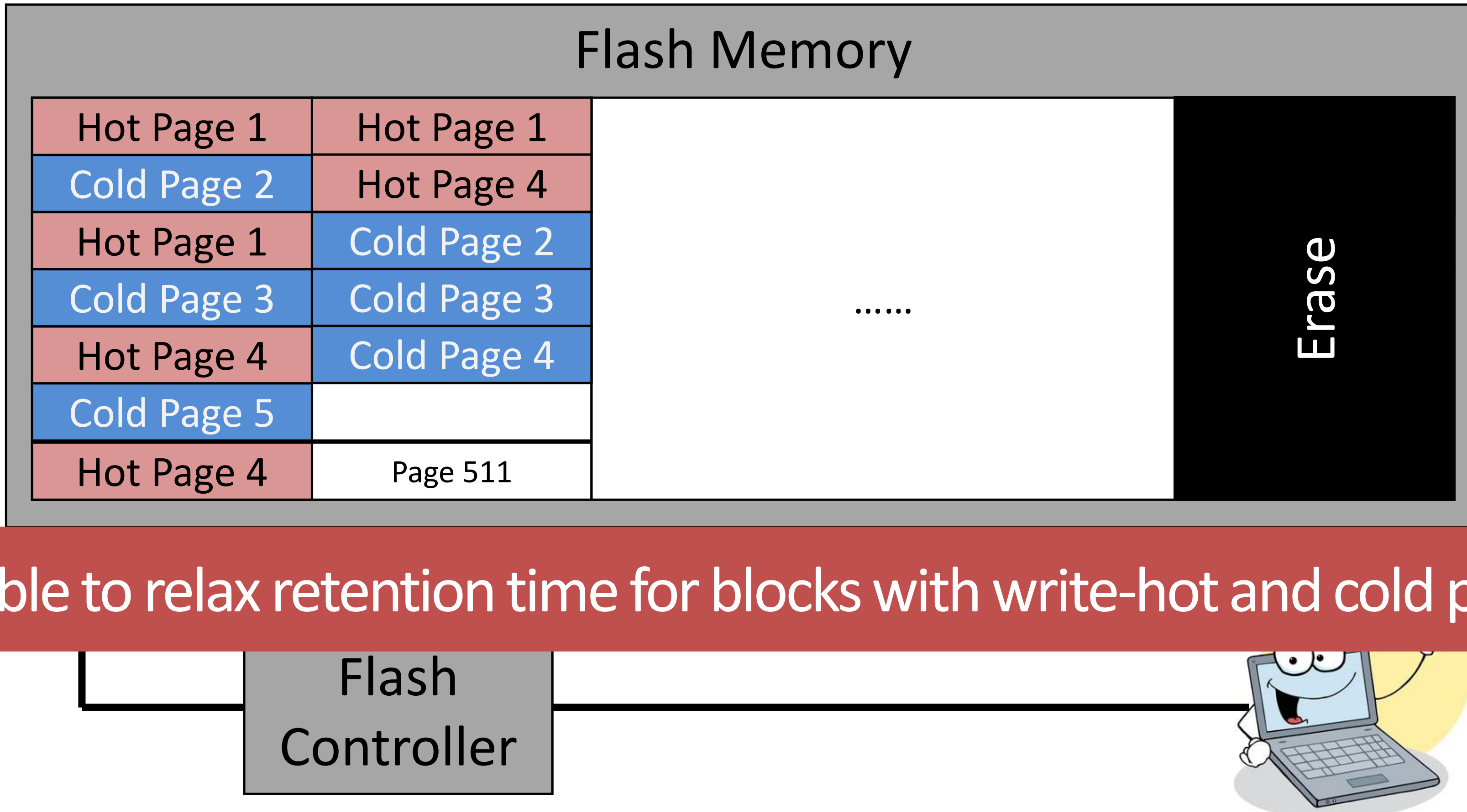
SAFARI

Carnegie Mellon

Executive Summary

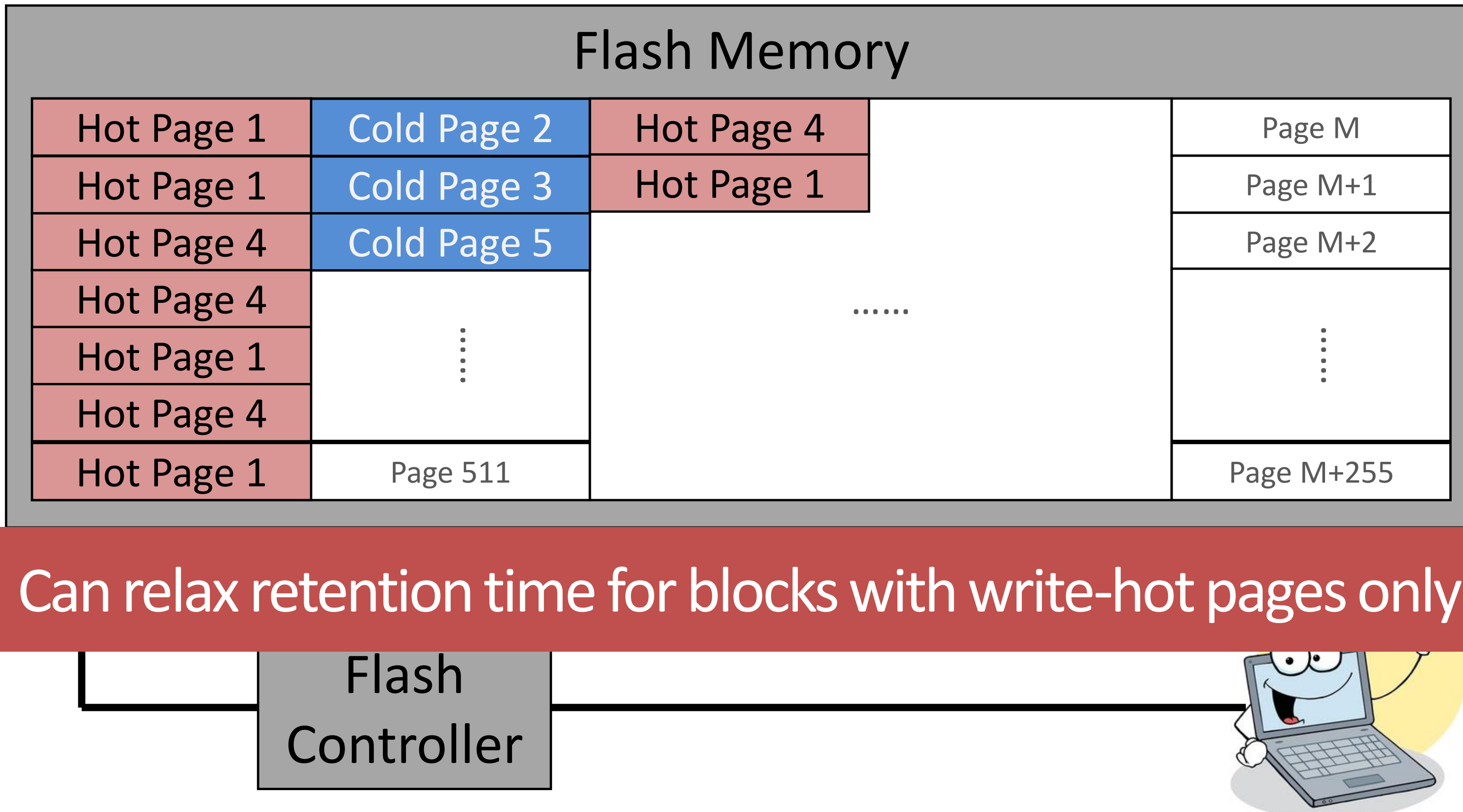
- Flash memory can achieve **50x endurance improvement by relaxing retention time using refresh** *[Cai+ ICCD '12]*
- *Problem*: **Frequent refresh consumes the majority of endurance improvement**
- *Goal*: Reduce refresh overhead to increase flash memory lifetime
- *Key Observation*: **Refresh is unnecessary for *write-hot data***
- *Key Ideas of Write-hotness Aware Retention Management (WARM)*
  - **Physically partition write-hot pages and write-cold pages** within the flash drive
  - **Apply *different* policies** (garbage collection, wear-leveling, refresh) to each group
- *Key Results*
  - WARM w/o refresh **improves lifetime by 3.24x**
  - WARM w/ adaptive refresh **improves lifetime by 12.9x** (1.21x over refresh only)

Conventional Write-Hotness Oblivious Management



Unable to relax retention time for blocks with write-hot and cold pages

Key Idea: Write-Hotness Aware Management



Can relax retention time for blocks with write-hot pages only

Write-Hotness Aware Retention Management

- Yixin Luo, Yu Cai, Saugata Ghose, Jongmoo Choi, and Onur Mutlu, **"WARM: Improving NAND Flash Memory Lifetime with Write-hotness Aware Retention Management"** *Proceedings of the 31st International Conference on Massive Storage Systems and Technologies (**MSST**)*, Santa Clara, CA, June 2015. [Slides (pptx) (pdf)] [Poster (pdf)]

WARM: Improving NAND Flash Memory Lifetime with Write-hotness Aware Retention Management

Yixin Luo
yixinluo@cs.cmu.edu

Yu Cai
yucaicai@gmail.com

Saugata Ghose
ghose@cmu.edu

Jongmoo Choi†
choijm@dankook.ac.kr

Onur Mutlu
onur@cmu.edu

Agenda

- Background, Motivation and Approach
- Experimental Characterization Methodology
- Error Analysis and Management
  - Main Characterization Results
  - Retention-Aware Error Management
  - Threshold Voltage and Program Interference Analysis
  - Read Reference Voltage Prediction
  - Neighbor-Assisted Error Correction
  - Read Disturb Error Handling
  - Retention Error Handling
  - Large Scale Field Analysis
  - 3D NAND Flash Memory Reliability
- Summary

Summary of Key Works

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu, **"Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives"** *Proceedings of the IEEE*, September 2017.

Cai+, "Error Patterns in MLC NAND Flash Memory: Measurement, Characterization, and Analysis," DATE 2012.
Cai+, "Flash Correct-and-Refresh: Retention-Aware Error Management for Increased Flash Memory Lifetime," ICCD 2012.
Cai+, "Threshold Voltage Distribution in MLC NAND Flash Memory: Characterization, Analysis and Modeling," DATE 2013.
Cai+, "Error Analysis and Retention-Aware Error Management for NAND Flash Memory," Intel Technology Journal 2013.
Cai+, "Program Interference in MLC NAND Flash Memory: Characterization, Modeling, and Mitigation," ICCD 2013.
Cai+, "Neighbor-Cell Assisted Error Correction for MLC NAND Flash Memories," SIGMETRICS 2014.
Cai+,"Data Retention in MLC NAND Flash Memory: Characterization, Optimization and Recovery," HPCA 2015.
Cai+, "Read Disturb Errors in MLC NAND Flash Memory: Characterization and Mitigation," DSN 2015.
Luo+, "WARM: Improving NAND Flash Memory Lifetime with Write-hotness Aware Retention Management," MSST 2015.
Meza+, "A Large-Scale Study of Flash Memory Errors in the Field," SIGMETRICS 2015.
Luo+, "Enabling Accurate and Practical Online Flash Channel Modeling for Modern MLC NAND Flash Memory," IEEE JSAC 2016.
Cai+, "Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques," HPCA 2017.
Fukami+, "Improving the Reliability of Chip-Off Forensic Analysis of NAND Flash Memory Devices," DFRWS EU 2017.
Luo+, "HeatWatch: Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature-Awareness," HPCA 2018.
Luo+, "Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation," SIGMETRICS 2018.

NAND Flash Vulnerabilities **[HPCA'17]**

HPCA, Feb. 2017

Vulnerabilities in MLC NAND Flash Memory Programming: Experimental Analysis, Exploits, and Mitigation Techniques

Yu Cai[†] Saugata Ghose[†] Yixin Luo[‡†] Ken Mai[†] Onur Mutlu[§†] Erich F. Haratsch[‡]
[†]*Carnegie Mellon University* [‡]*Seagate Technology* [§]*ETH Zürich*

Modern NAND flash memory chips provide high density by storing two bits of data in each flash cell, called a multi-level cell *(MLC). An MLC partitions the threshold voltage range of a flash cell into four* voltage states. *When a flash cell is programmed, a high voltage is applied to the cell. Due to parasitic capacitance coupling between flash cells that are physically close to each other, flash cell programming can lead to* cell-to-cell program interference, *which introduces errors into* neighboring *flash cells. In order to reduce the impact of cell-to-cell interference on the reliability of MLC NAND flash memory, flash manufacturers adopt a* two-step programming method, *which programs the MLC in two separate steps. First, the flash memory partially programs the least significant bit of the MLC to some intermediate threshold voltage. Second, it programs the most significant bit to bring the MLC up to its full voltage state.*

In this paper, we demonstrate that two-step programming exposes new reliability and security vulnerabilities. We expe-

belongs to a different flash memory *page* (the unit of data programmed and read at the same time), which we refer to, respectively, as the least significant bit (LSB) page and the most significant bit (MSB) page [5].

A flash cell is programmed by applying a large voltage on the control gate of the transistor, which triggers charge transfer into the floating gate, thereby increasing the threshold voltage. To precisely control the threshold voltage of the cell, the flash memory uses *incremental step pulse programming* (ISPP) [12, 21, 25, 41]. ISPP applies multiple short pulses of the programming voltage to the control gate, in order to increase the cell threshold voltage by some small voltage amount (V_{step}) after each step. Initial MLC designs programmed the threshold voltage in *one shot*, issuing all of the pulses back-to-back to program *both* bits of data at the same time. However, as flash memory scales down, the distance between neighboring flash cells decreases, which

https://people.inf.ethz.ch/omutlu/pub/flash-memory-programming-vulnerabilities_hpca17.pdf

NAND Flash Errors: A Modern Survey

INVITED PAPER

Proceedings of the IEEE, Sept. 2017

Error Characterization, Mitigation, and Recovery in Flash-Memory-Based Solid-State Drives

This paper reviews the most recent advances in solid-state drive (SSD) error characterization, mitigation, and data recovery techniques to improve both SSD's reliability and lifetime.

By Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu

https://arxiv.org/pdf/1706.08642

More Up-to-date Version

- Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu, **"Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery"**
Invited Book Chapter in Inside Solid State Drives, 2018.
[Preliminary arxiv.org version]

Errors in Flash-Memory-Based Solid-State Drives: Analysis, Mitigation, and Recovery

YU CAI, SAUGATA GHOSE
Carnegie Mellon University

ERICH F. HARATSCH
Seagate Technology

YIXIN LUO
Carnegie Mellon University

ONUR MUTLU
ETH Zürich and Carnegie Mellon University

Computer Architecture

Lecture 15a: Flash Memory and Solid-State Drives (II)

Prof. Onur Mutlu

ETH Zürich

Fall 2018

14 November 2018

Other Works on Flash Memory

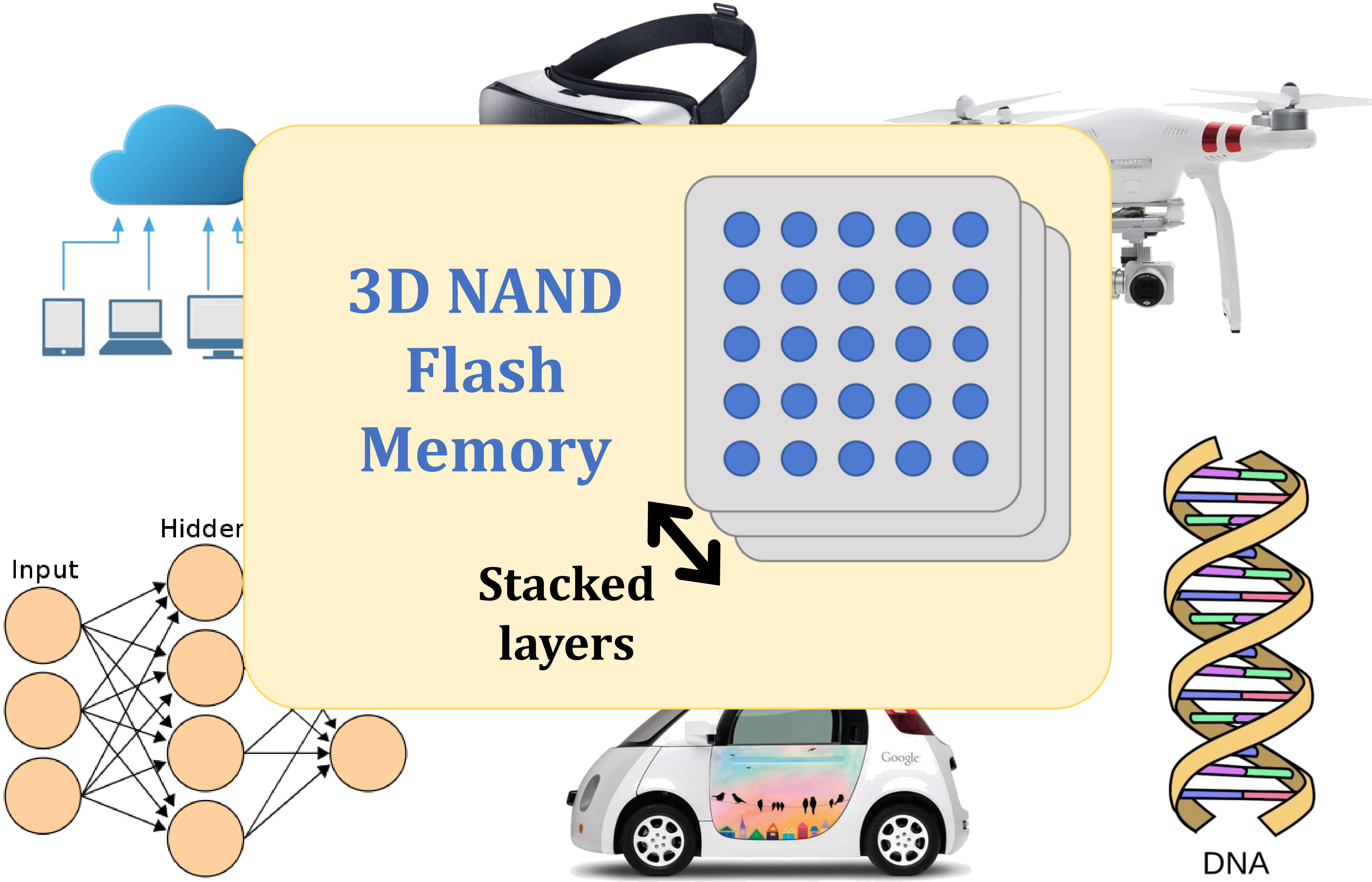
HeatWatch

Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature Awareness

Yixin Luo **Saugata Ghose** **Yu Cai** **Erich F. Haratsch** **Onur Mutlu**

Carnegie Mellon
SK hynix
ETH Zürich
SAFARI
SEAGATE

Storage Technology Drivers - 2018



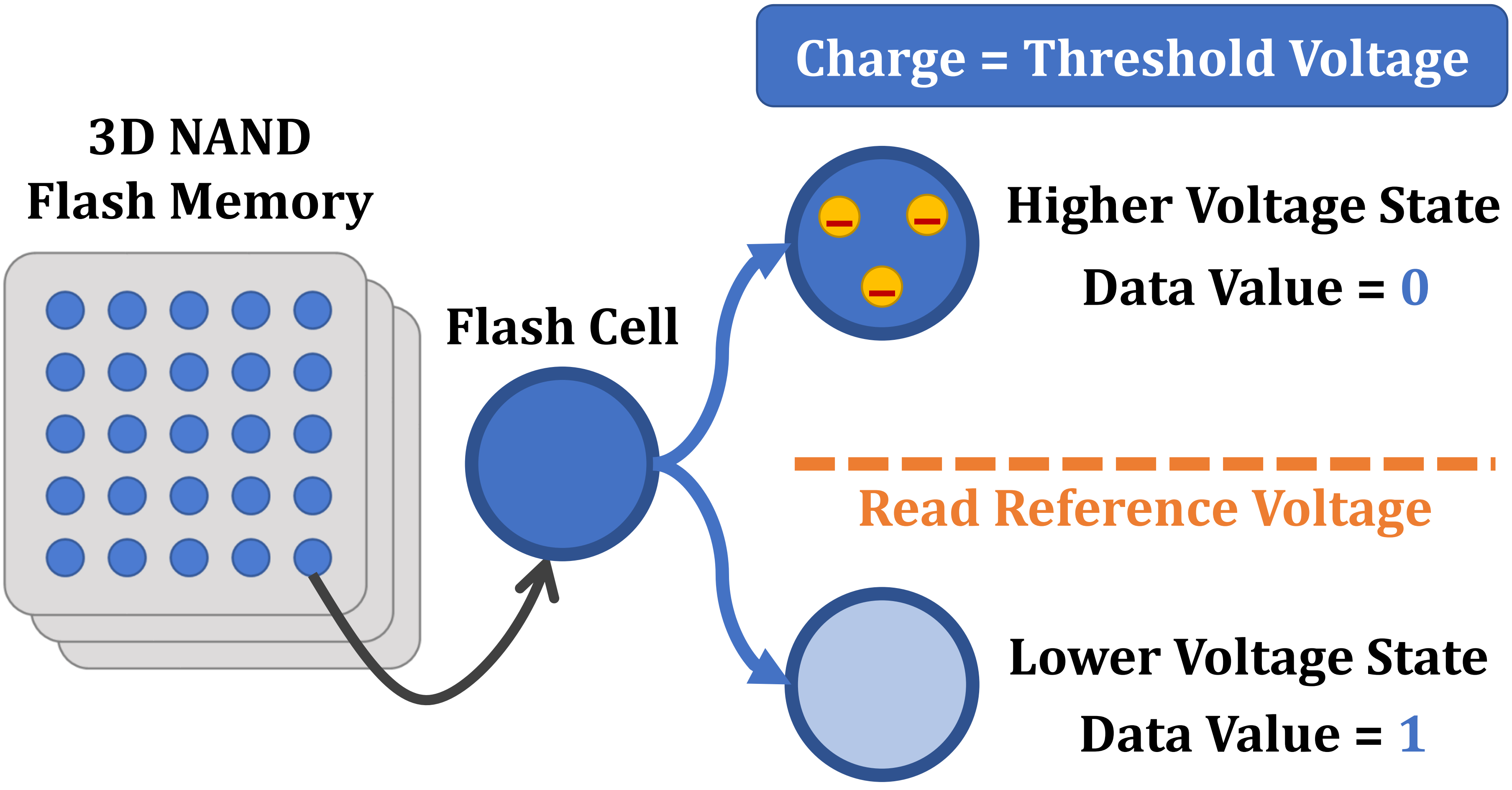
Executive Summary

- 3D NAND flash memory susceptible to **retention errors**
  - Charge leaks out of flash cell
  - Two unreported factors: ***self-recovery* and *temperature***
- We study *self-recovery* and *temperature* effects
  - **Experimental characterization** of *real* 3D NAND chips
  - **Unified Self-Recovery and Temperature (URT) Model**
    - Predicts impact of retention loss, wearout, self-recovery, temperature on **flash cell voltage**
    - **Low prediction error rate: 4.9%**
- We develop a new technique to improve flash reliability
  - **HeatWatch**
    - Uses URT model to find optimal read voltages for 3D NAND flash
    - **Improves flash lifetime by 3.85x**

Outline

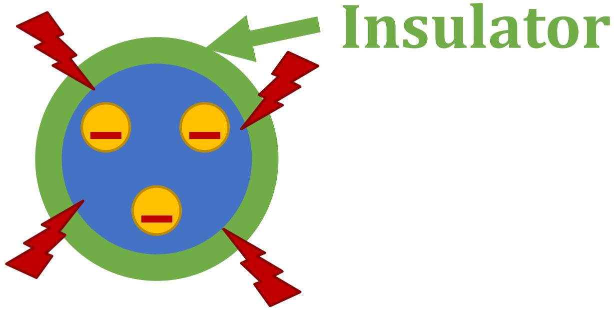
- Executive Summary
- **Background on NAND Flash Reliability**
- Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips
- URT: Unified Self-Recovery and Temperature Model
- HeatWatch Mechanism
- Conclusion

3D NAND Flash Memory Background



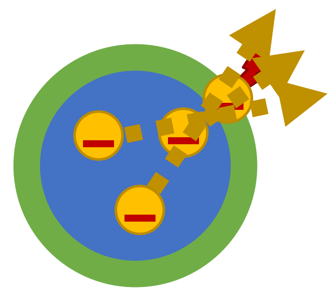
Flash Wearout

Program/Erase (P/E) → Wearout

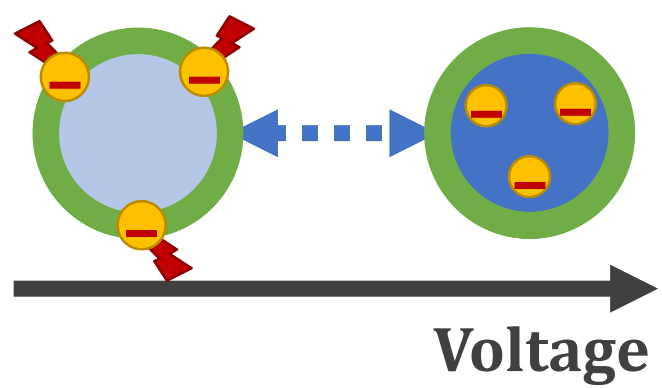


Wearout Effects:

1. Retention Loss
(voltage shift over time)



2. Program Variation
(init. voltage difference b/w states)



Wearout Introduces Errors

Improving Flash Lifetime

Errors introduced by wearout
limit flash lifetime
(measured in P/E cycles)

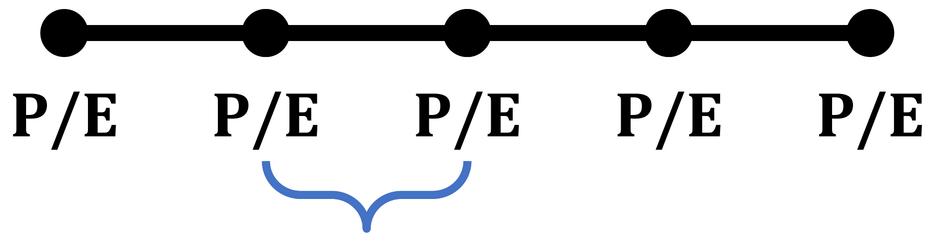
Two Ways to Improve Flash Lifetime ➡

Exploiting the Self-Recovery Effect

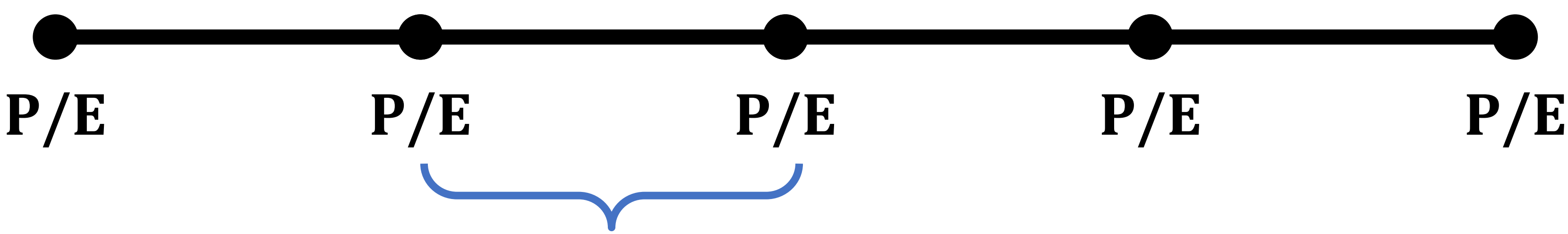
Exploiting the Temperature Effect

Exploiting the Self-Recovery Effect

Partially repairs damage due to wearout



Dwell Time: Idle Time Between P/E Cycles

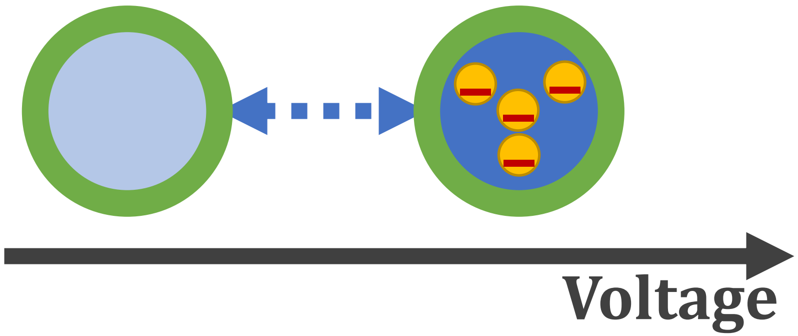


Longer Dwell Time: More Self-Recovery

Reduces Retention Loss

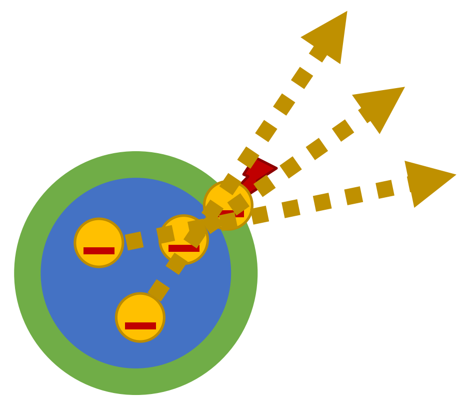
Exploiting the Temperature Effect

High Program Temperature



Increases Program Variation

High Storage Temperature



Accelerates Retention Loss

Prior Studies of Self-Recovery/Temperature

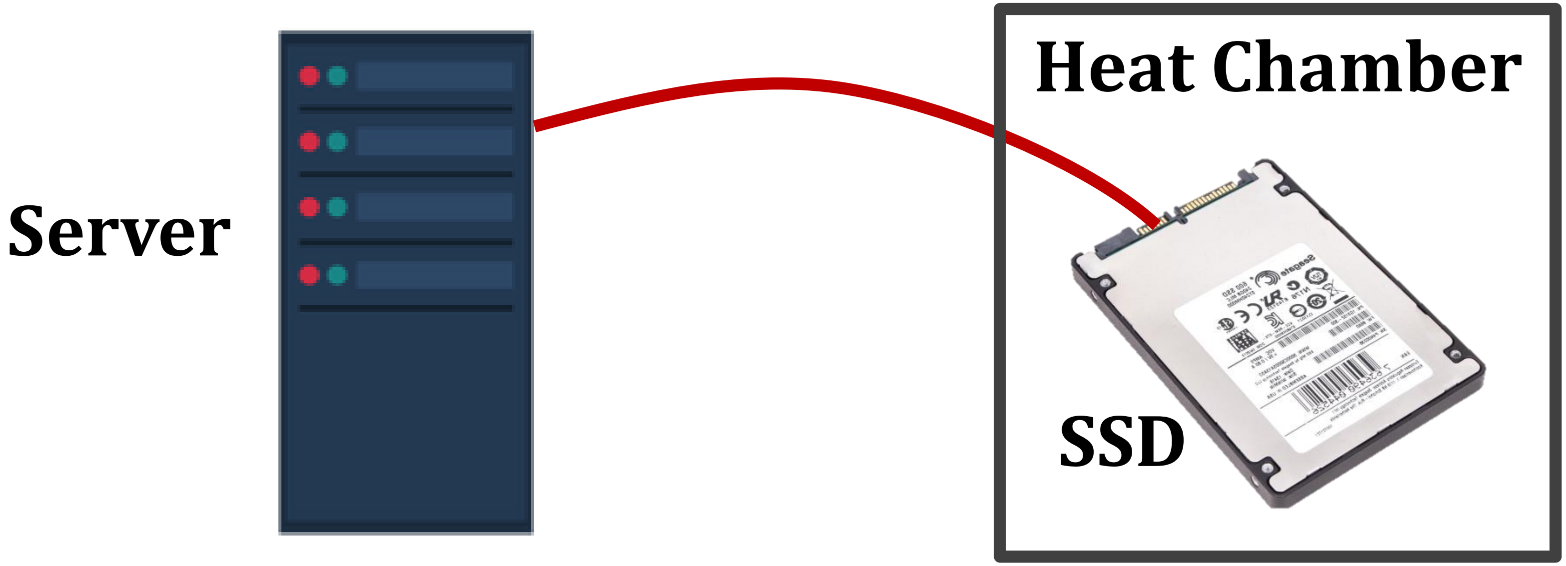
	Planar (2D) NAND	3D NAND
Self-Recovery Effect	✓ Mielke 2006	✗
Temperature Effect	✓ JEDEC 2010 (no characterization)	✗

Outline

- Executive Summary
- Background on NAND Flash Reliability
- **Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips**
- URT: Unified Self-Recovery and Temperature Model
- HeatWatch Mechanism
- Conclusion

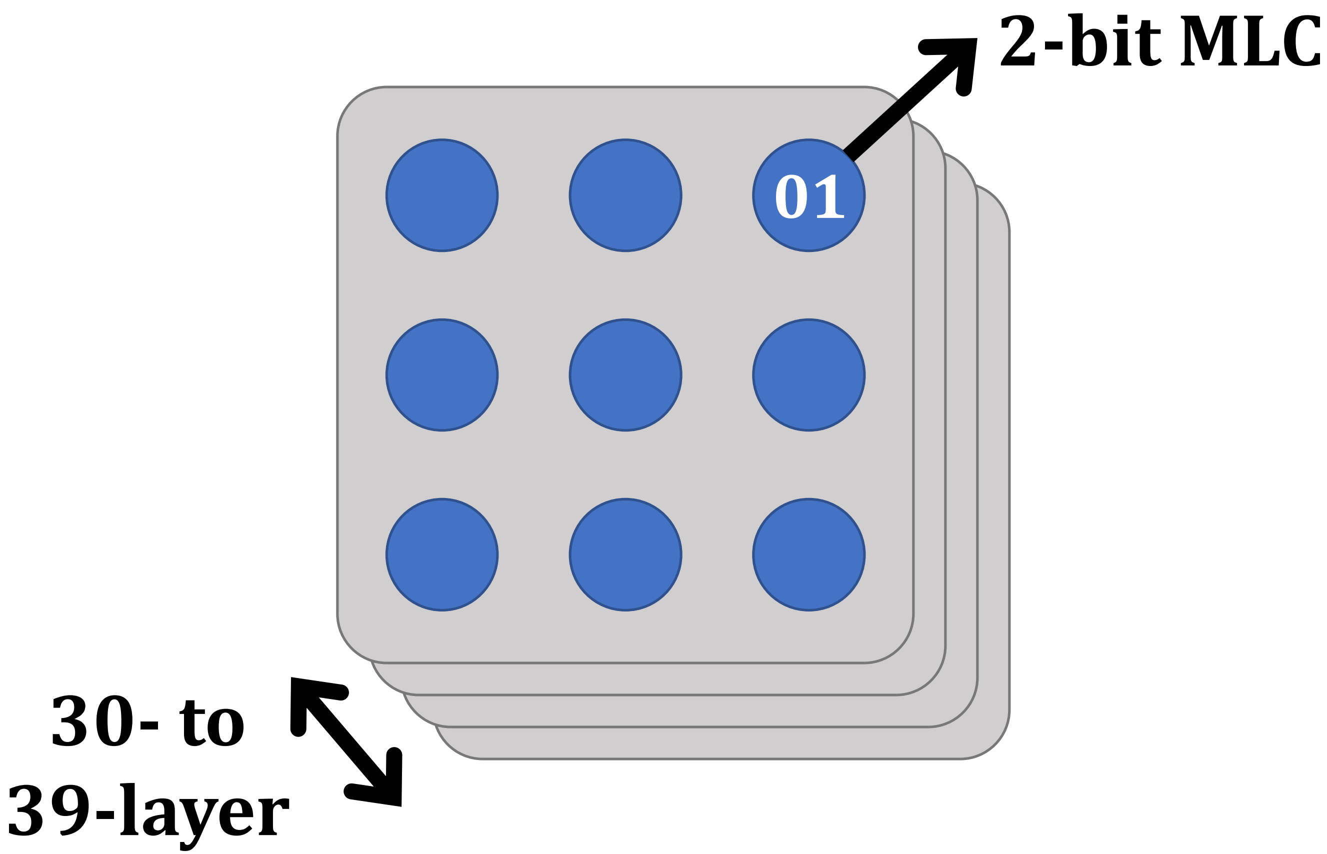
Characterization Methodology

- Modified firmware version in the flash controller
  - Control the read reference voltage of the flash chip
  - Bypass ECC to get raw NAND data (with raw bit errors)
- Control temperature with a heat chamber

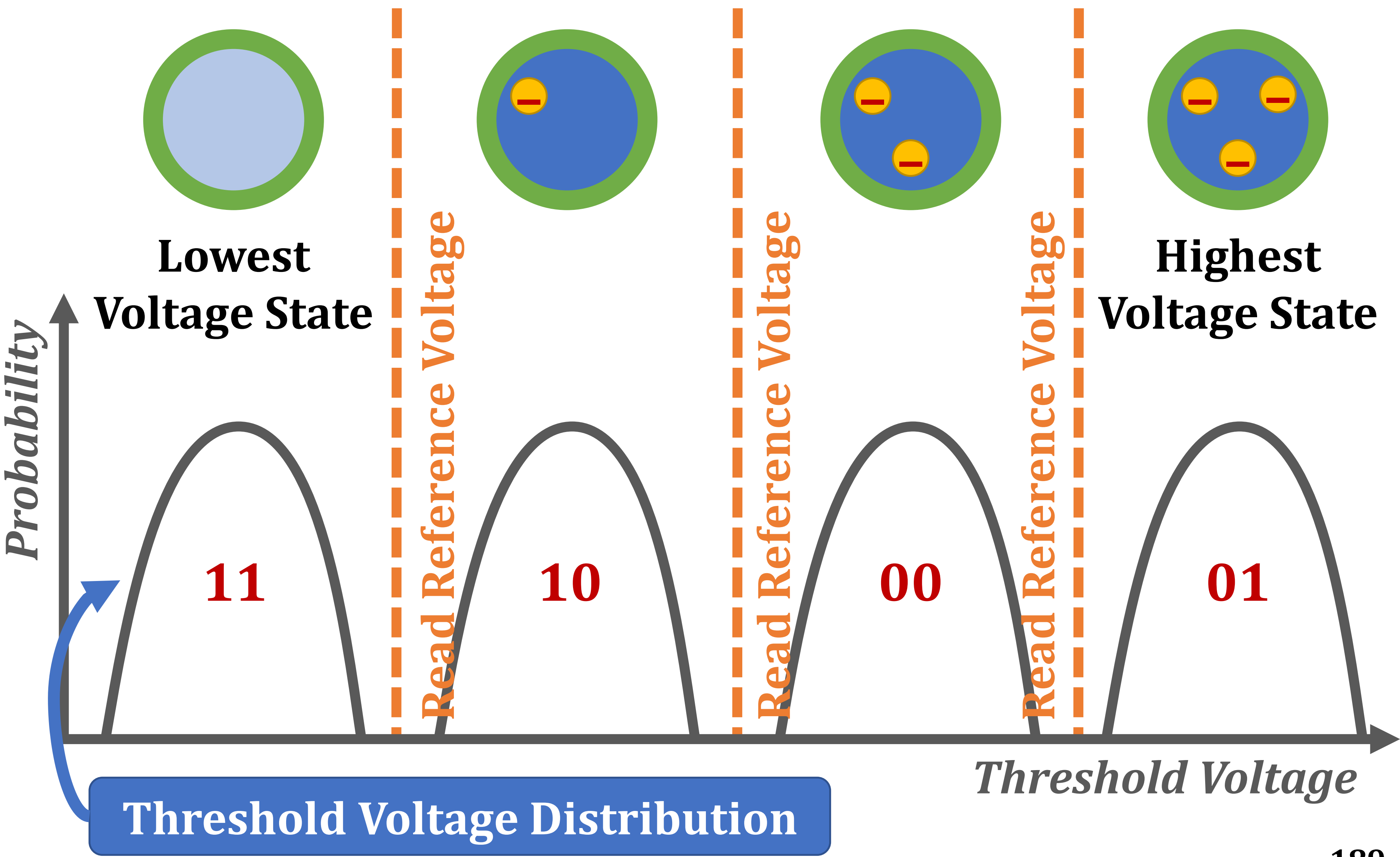


Characterized Devices

Real 30-39 Layer 3D MLC NAND Flash Chips

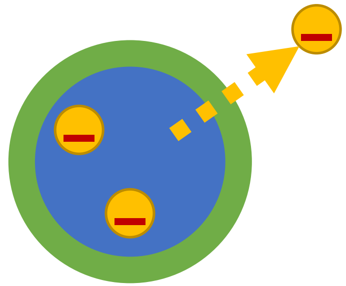


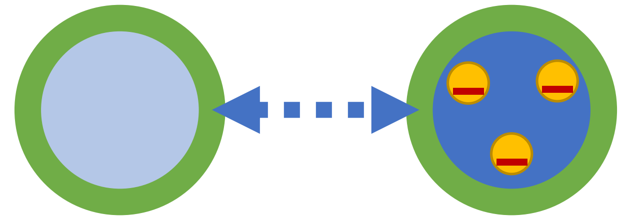
MLC Threshold Voltage Distribution Background



Characterization Goal

Characterized Metrics





Retention Loss Speed
(how fast voltage shifts over time)

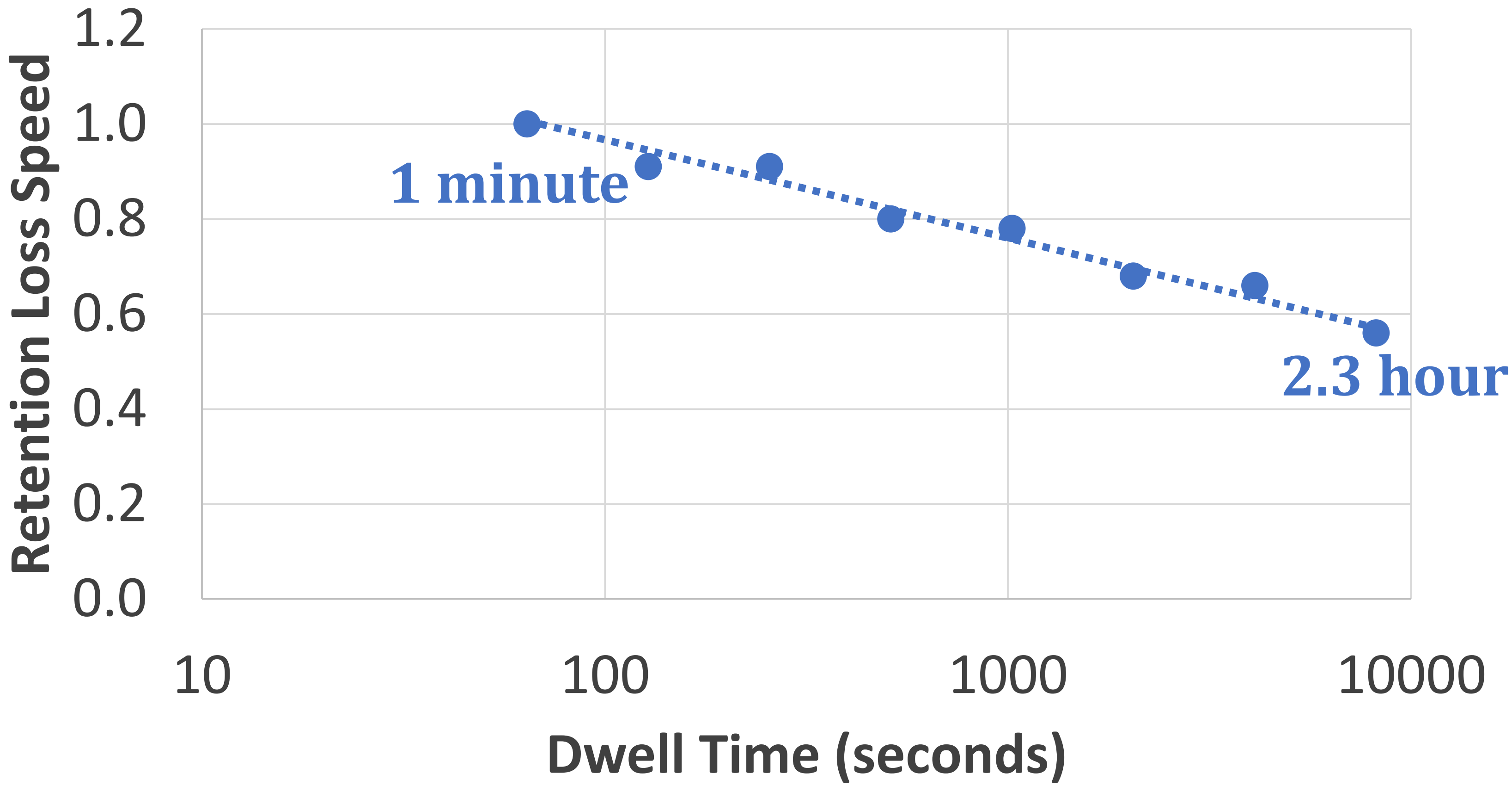
Program Variation
(initial voltage difference between states)

Characterized Phenomena

Self-Recovery Effect

Temperature Effect

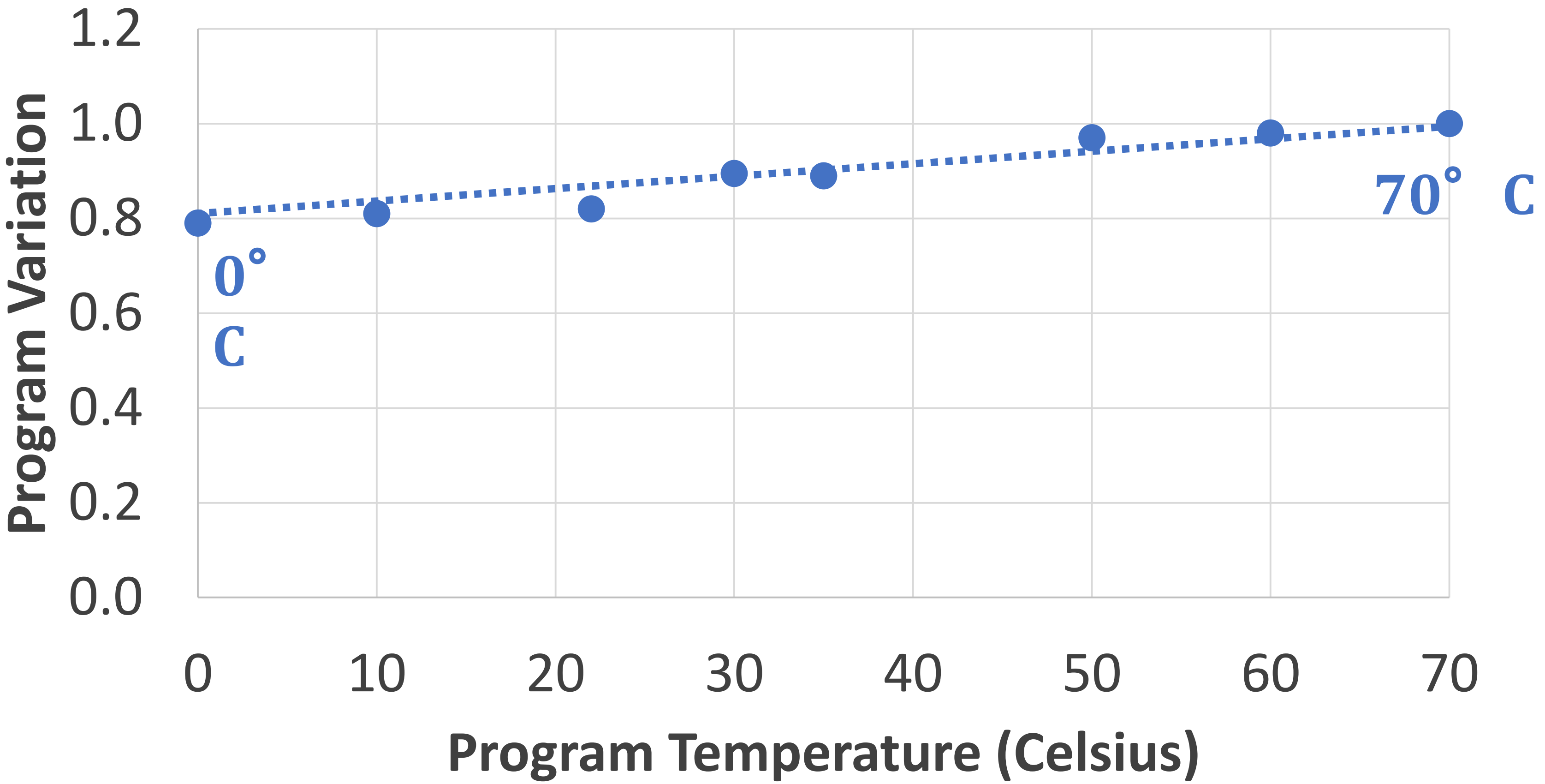
Self-Recovery Effect Characterization Results



Dwell time: Idle time between P/E cycles

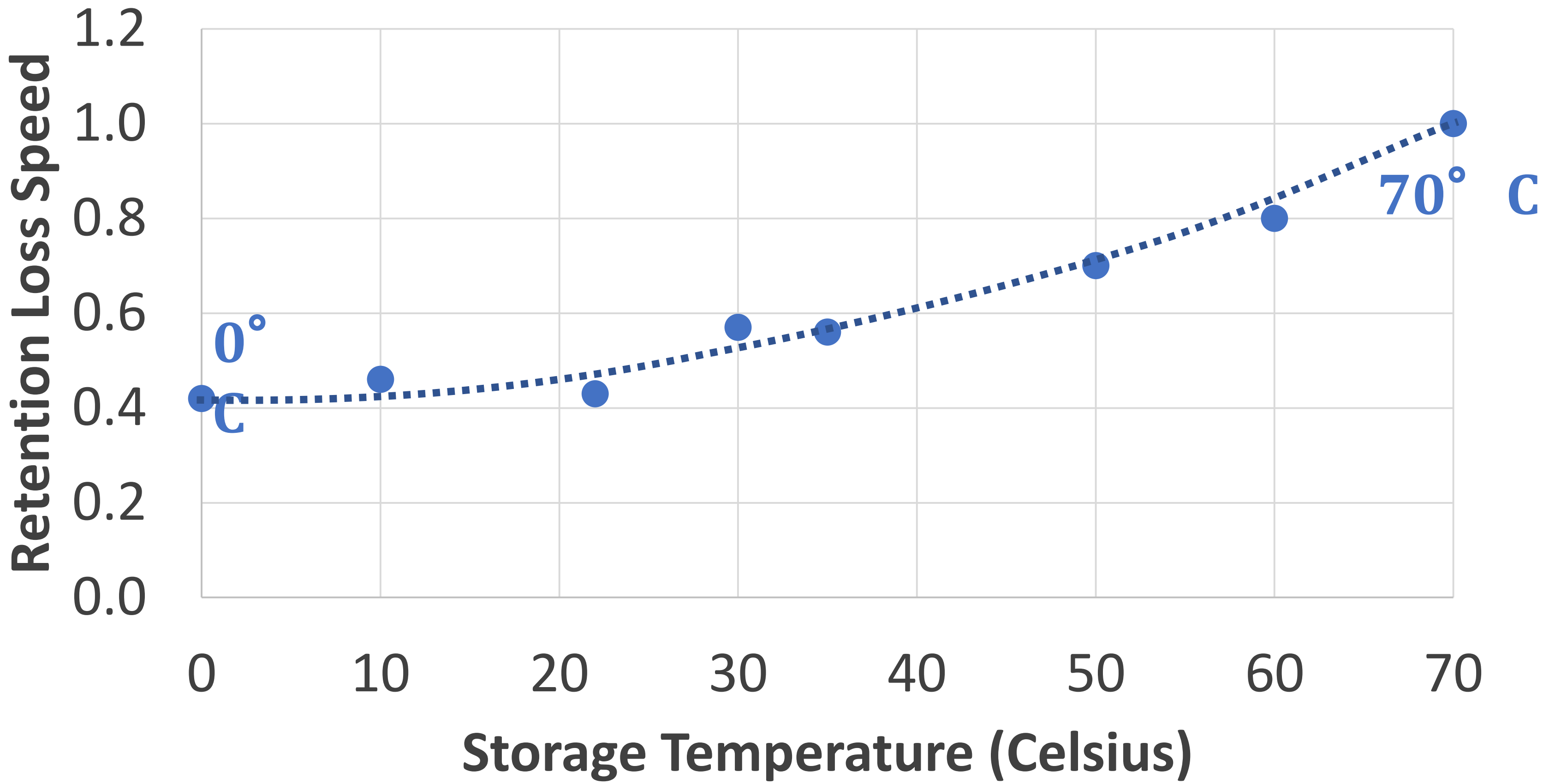
Increasing dwell time from 1 minute to 2.3 hours slows down retention loss speed by 40%

Program Temperature Effect Characterization Results



Increasing program temperature from 0°C to 70°C improves program variation by 21%

Storage Temperature Effect Characterization Results



Lowering storage temperature from 70°C to 0°C slows down retention loss speed by 58%

Characterization Summary

Major Results:

- *Self-recovery* affects retention loss speed
- Program *temperature* affects program variation
- *Storage temperature* affects retention loss speed

Unified Model

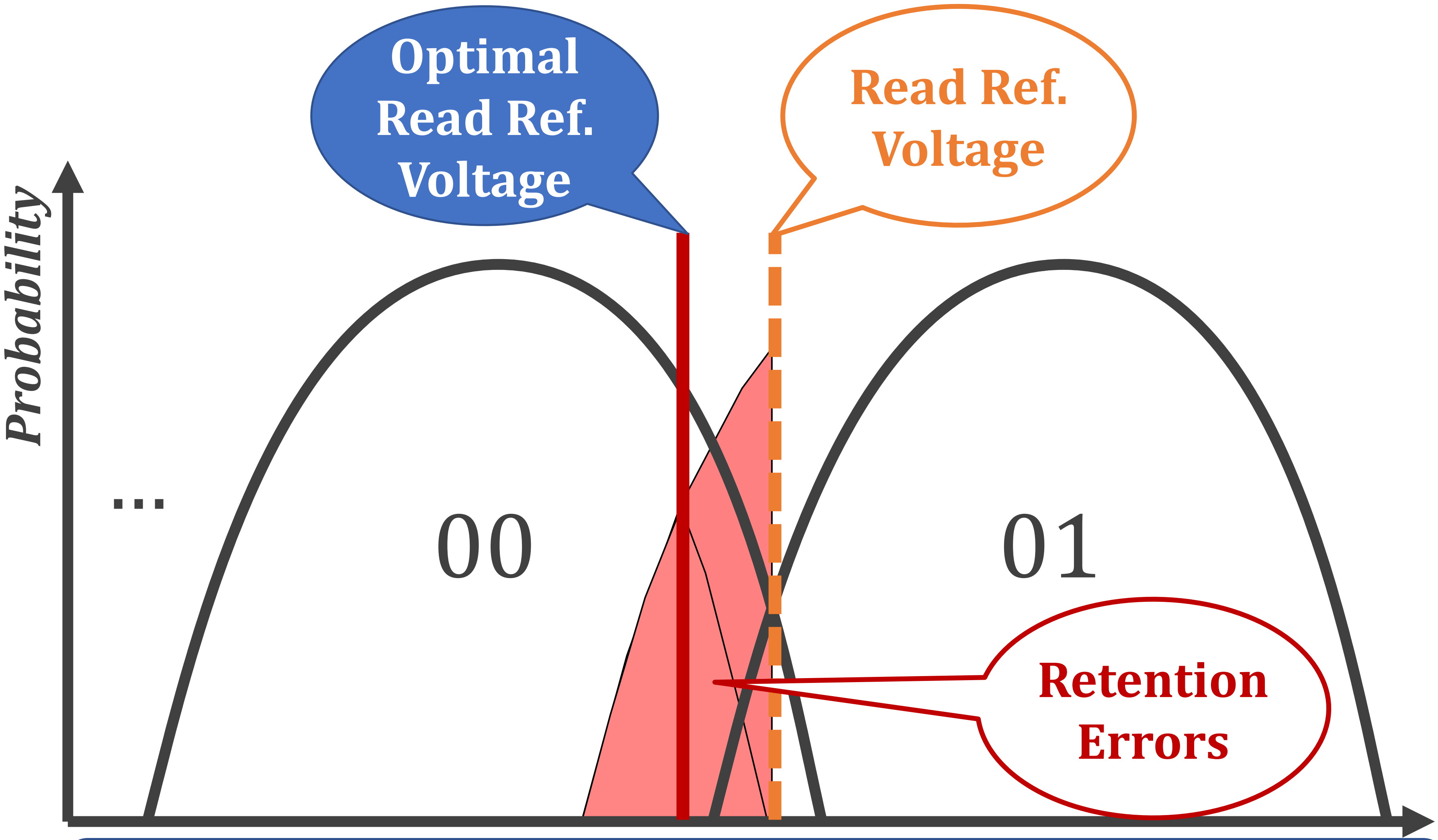
Other Characterizations Methods in the Paper:

- More detailed results on self-recovery and temperature
  - Effects on error rate
  - Effects on threshold voltage distribution
- Effects of recovery cycle (P/E cycles with long dwell time) on retention loss speed

Outline

- Executive Summary
- Background on NAND Flash Reliability
- Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips
- **URT: Unified Self-Recovery and Temperature Model**
- HeatWatch Mechanism
- Conclusion

Minimizing 3D NAND Errors



Optimal read reference voltage minimizes 3D NAND errors

Predicting the Mean Threshold Voltage

Our URT Model:

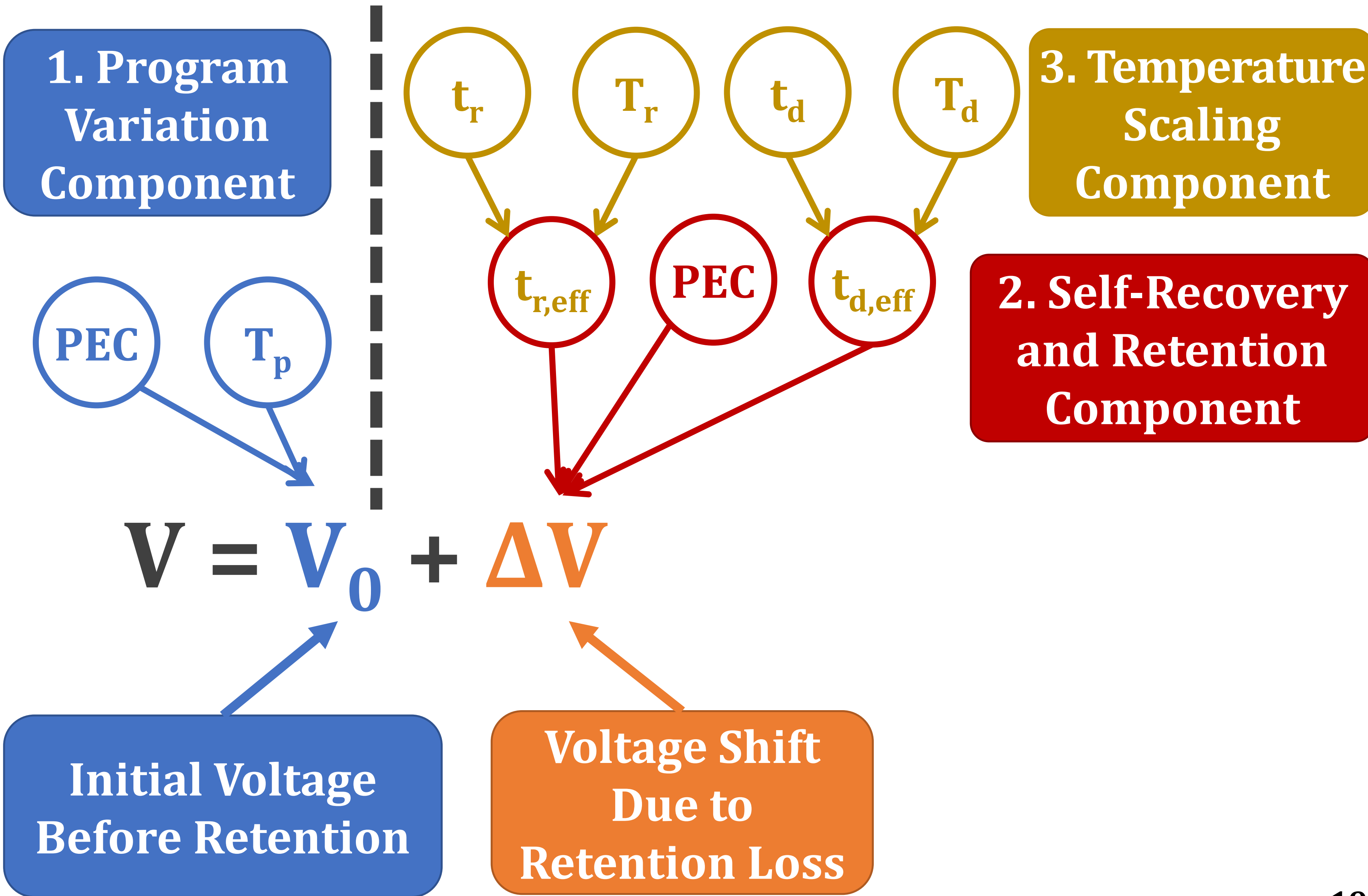
$$V = V_0 + \Delta V$$

Mean Threshold Voltage

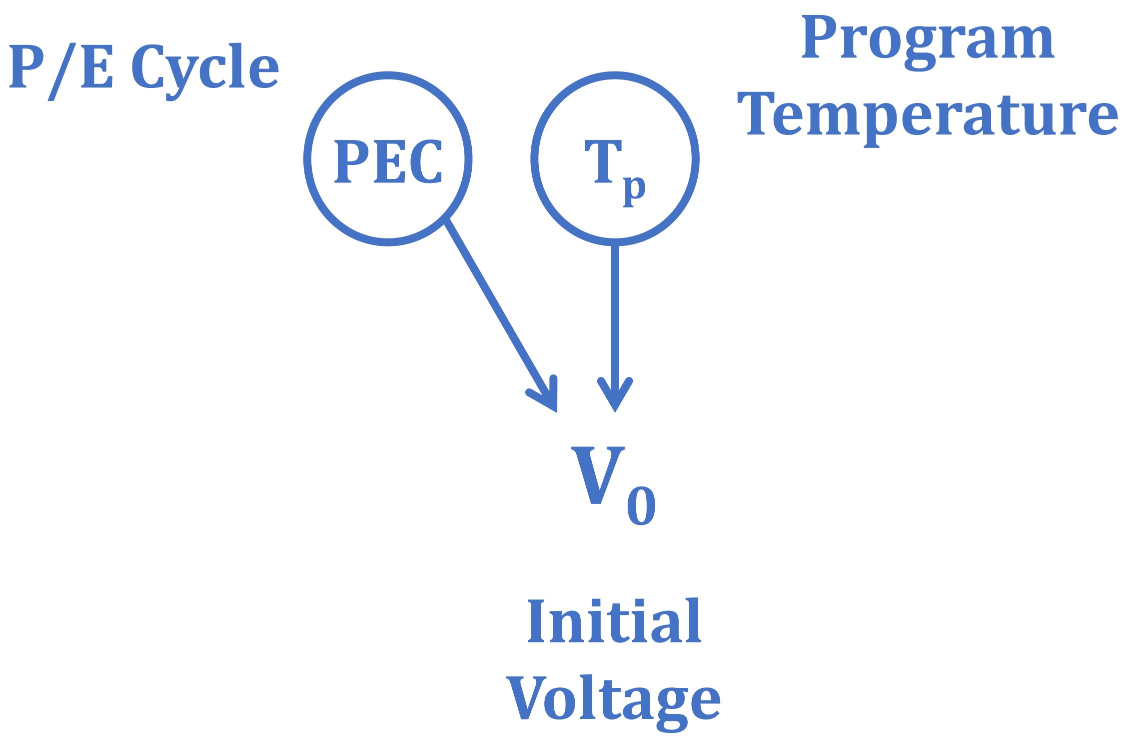
Initial Voltage Before Retention (Program Variation)

Voltage Shift Due to Retention Loss

URT Model Overview



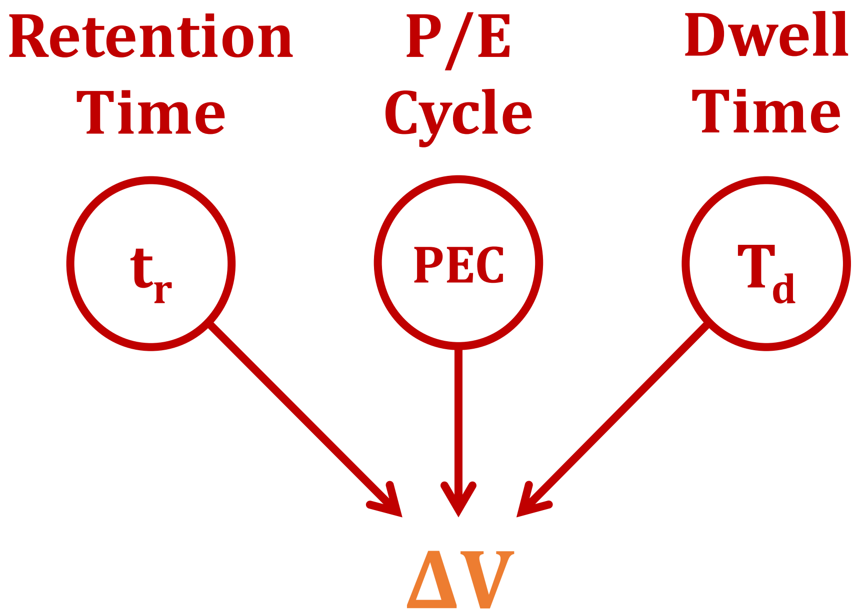
1. Program Variation Component



$$V_0 = A \cdot T_p \cdot PEC + B \cdot T_p + C \cdot PEC + D$$

Validation: R^2 = 91.7%

2. Self-Recovery and Retention Component

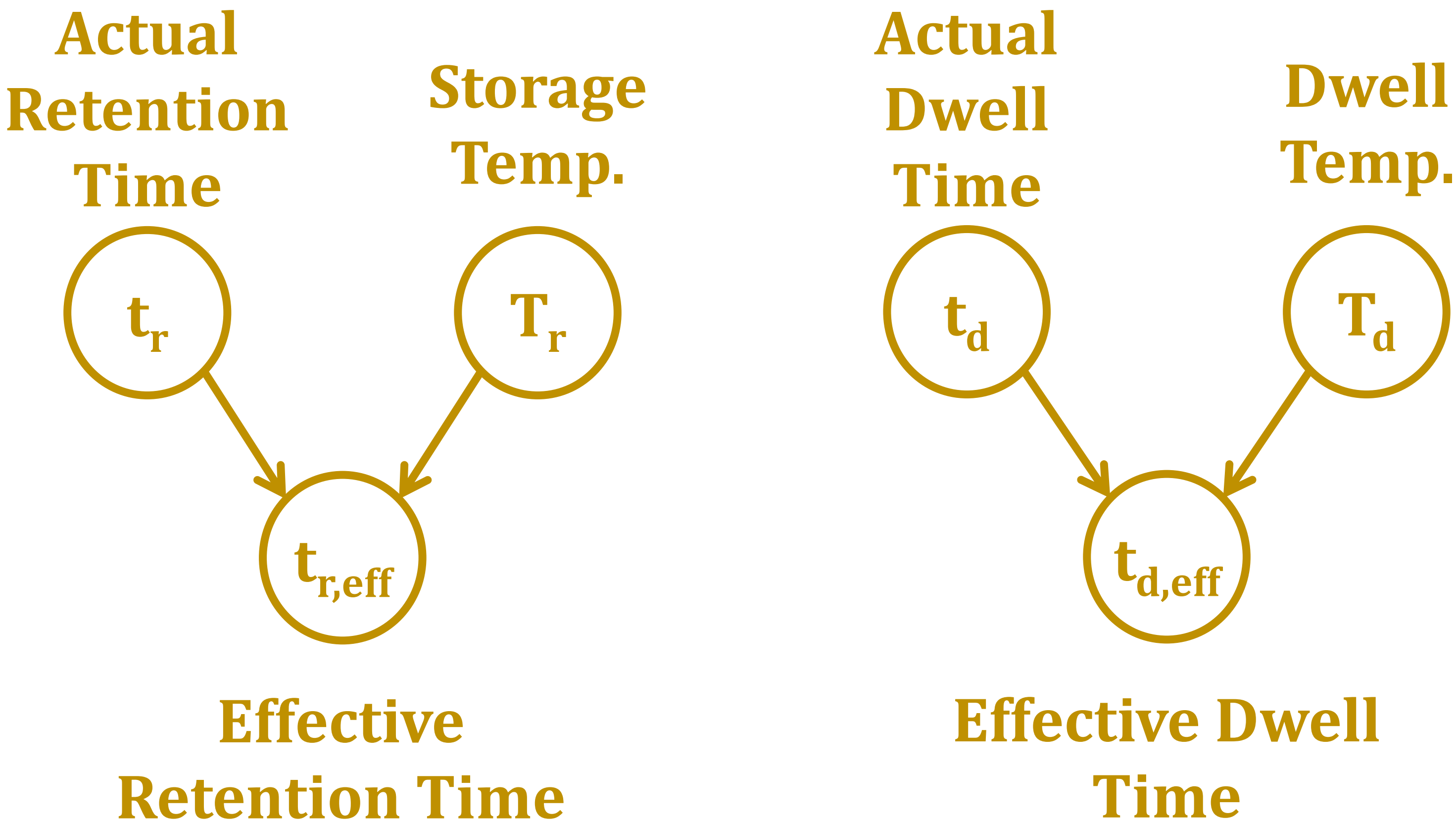


Retention Shift

$$\Delta V\,(t_{er}, t_{ed}, PEC) = b \cdot (PEC + c) \cdot \ln\left(1 + \frac{t_{er}}{t_0 + a \cdot t_{ed}}\right)$$

Validation: 3x more accurate than state-of-the-art model

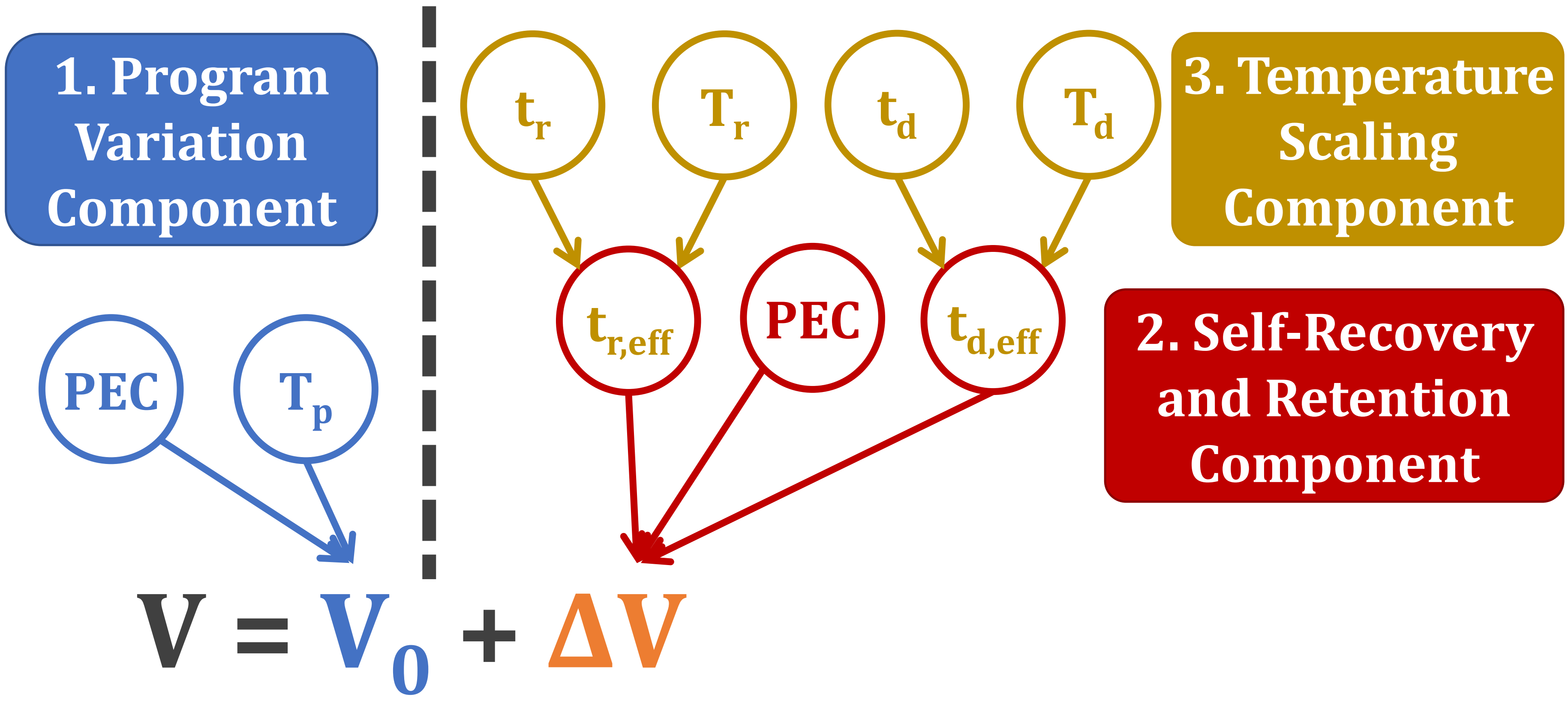
3. Temperature Scaling Component



Arrhenius Equation: $AF = \frac{t_{real}}{t_{room}} = \exp\left(\frac{E_a}{k_B} \cdot \left(\frac{1}{T_{real}} - \frac{1}{T_{room}}\right)\right)$

Validation: Adjust an important parameter, E_a, from 1.1 eV to 1.04 eV

URT Model Summary



$$V = V_0 + \Delta V$$

Validation:
Prediction Error Rate = 4.9%

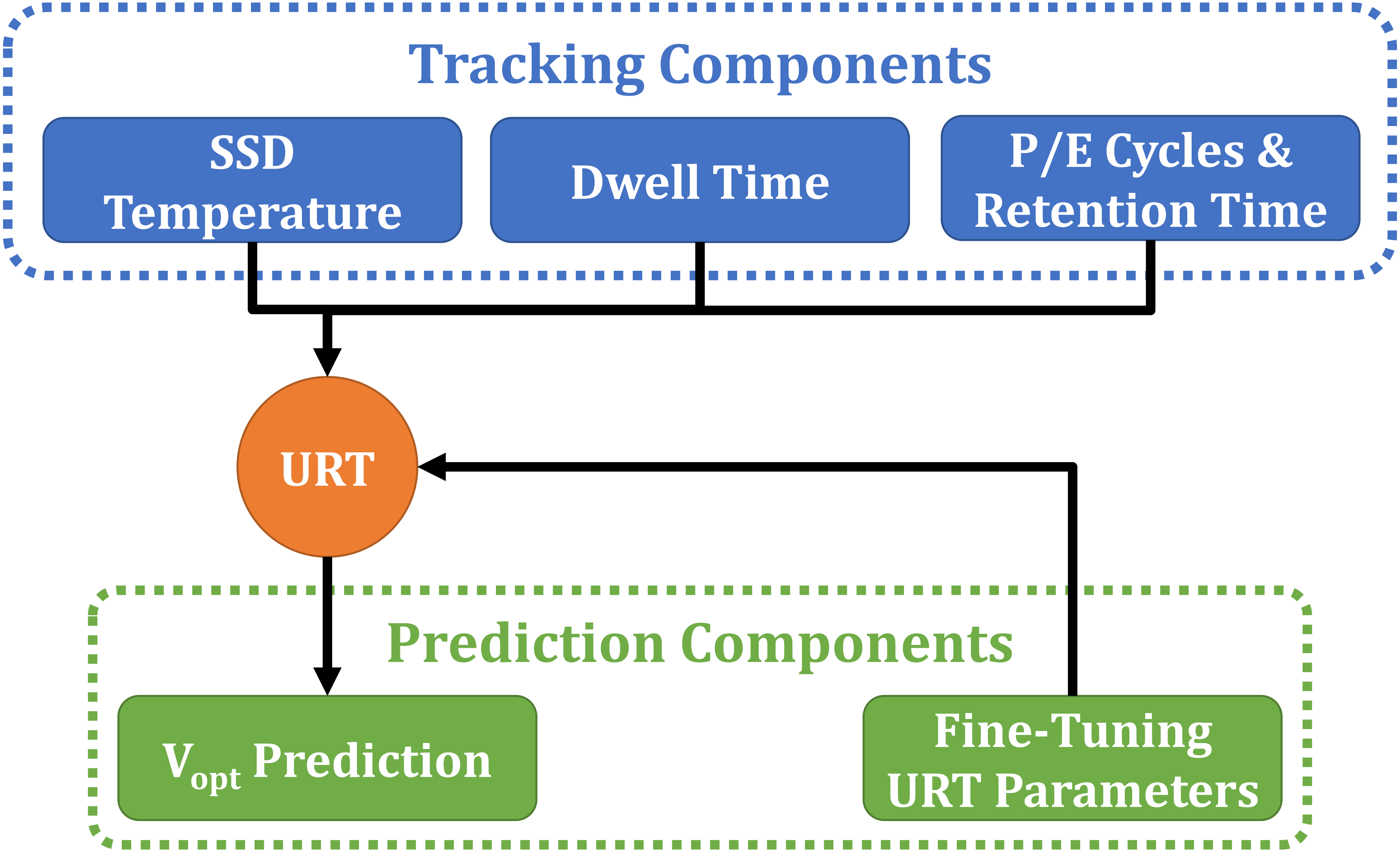
Outline

- Executive Summary
- Background on NAND Flash Reliability
- Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips
- URT: Unified Self-Recovery and Temperature Model
- **HeatWatch Mechanism**
- Conclusion

HeatWatch Mechanism

- **Key Idea**
  - **Predict change in threshold voltage distribution** by using the URT model
  - **Adapt read reference voltage to near-optimal** (V_{opt}) based on predicted change in voltage distribution

HeatWatch Mechanism Overview



Tracking SSD Temperature

Tracking Components

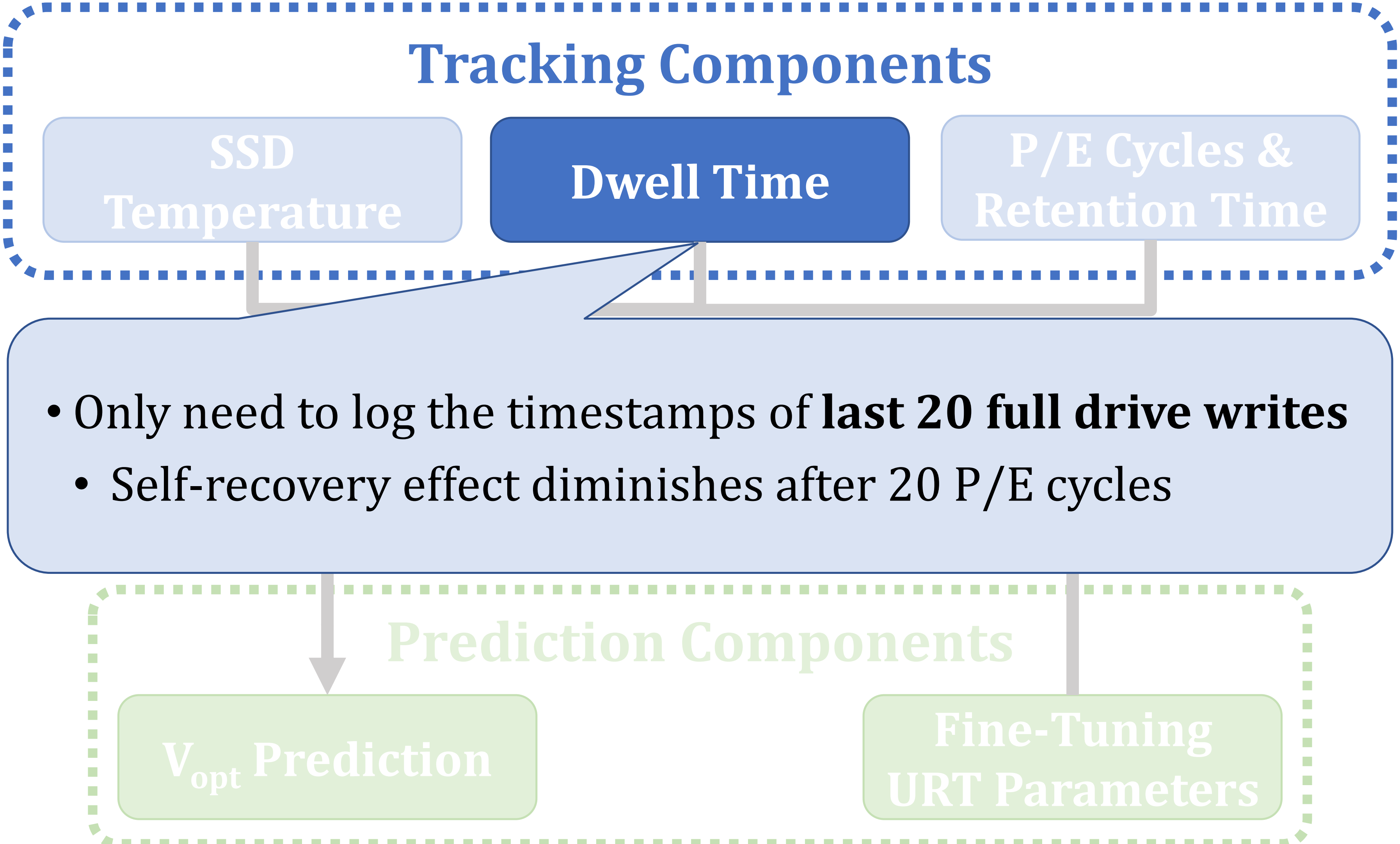
SSD Temperature | Dwell Time | P/E Cycles & Retention Time

- Use existing sensors in the SSD
- **Precompute** temperature scaling factor at **logarithmic time intervals**

Prediction Components

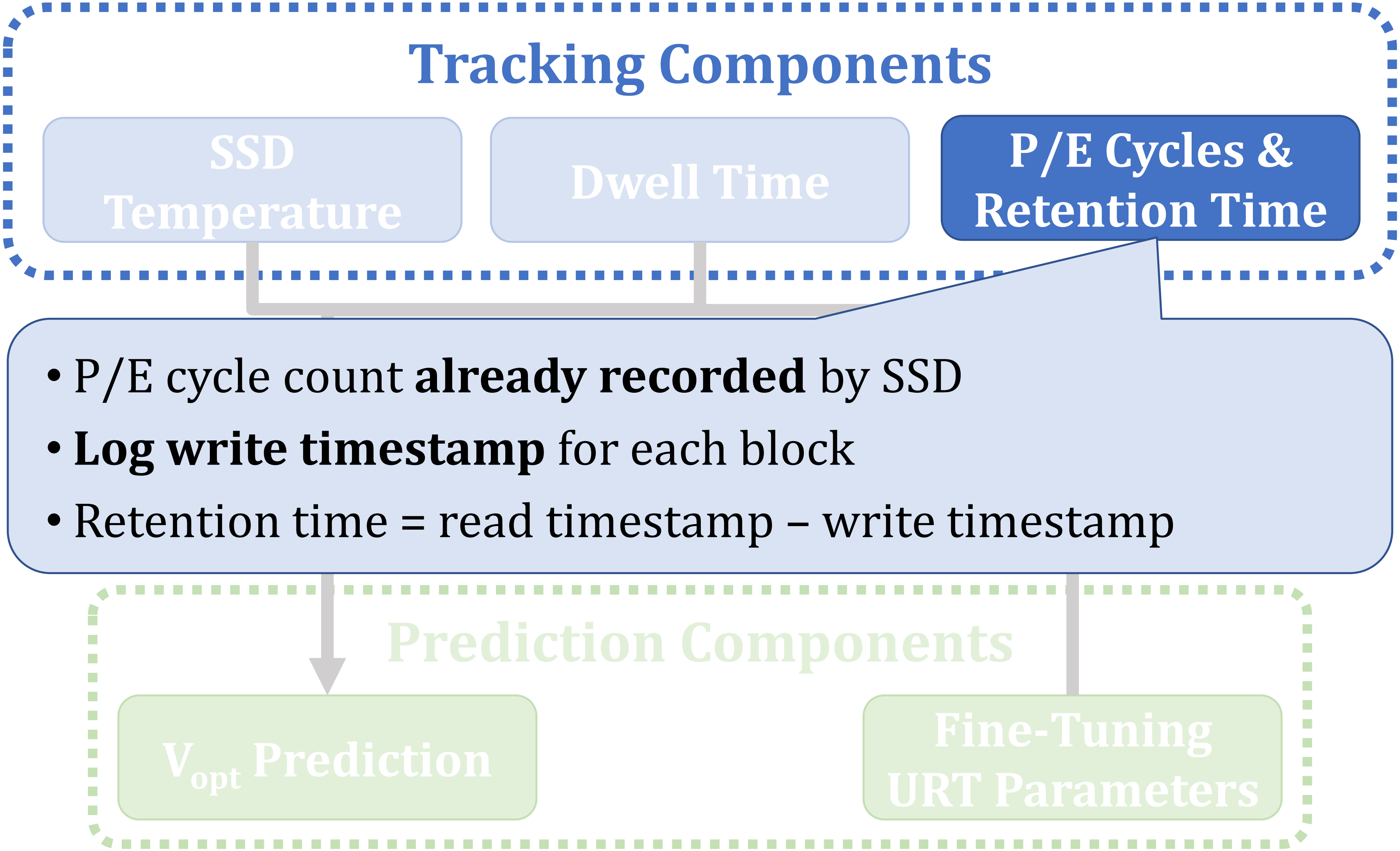
V_{opt} Prediction | Fine-Tuning URT Parameters

Tracking Dwell Time

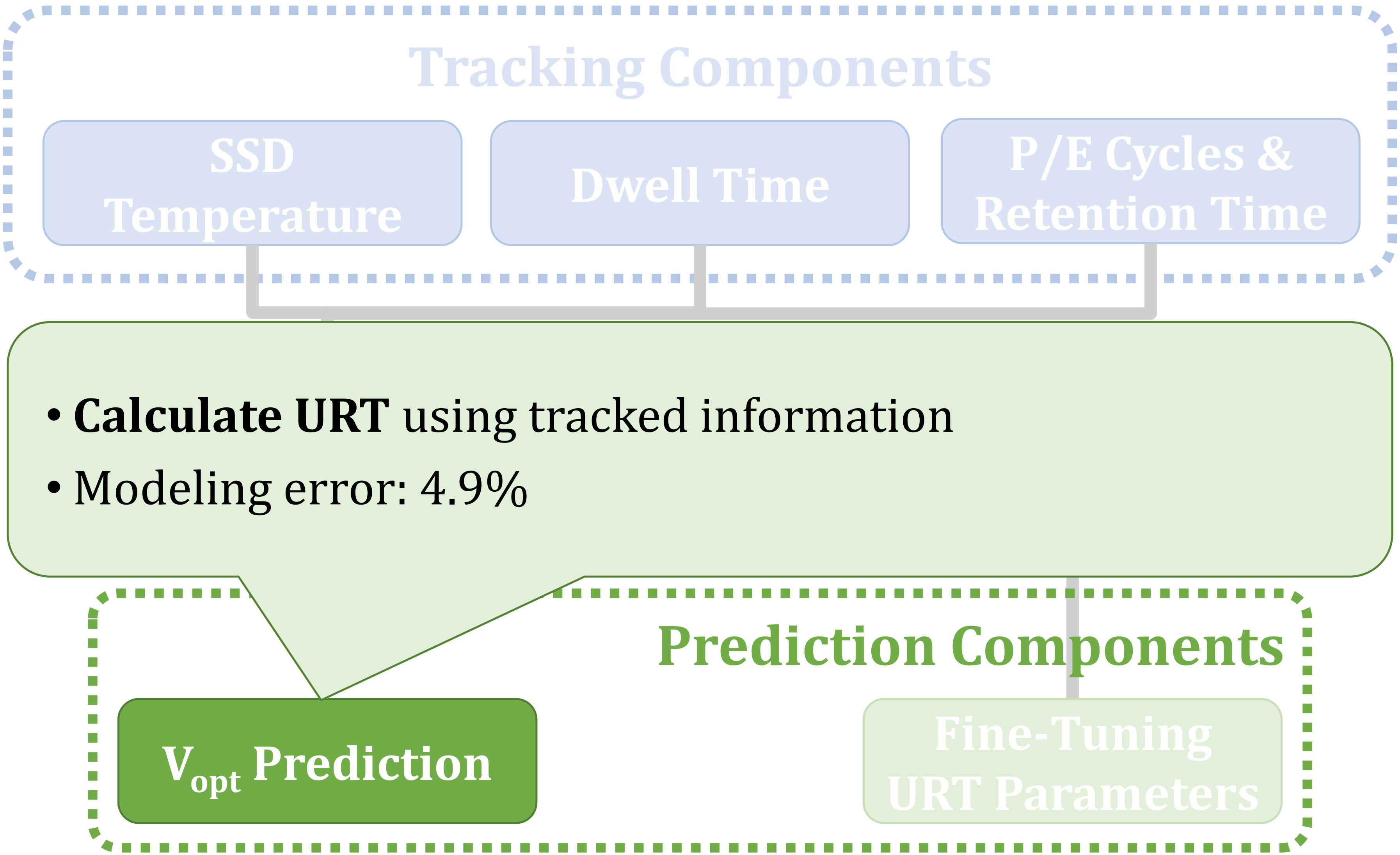


- Only need to log the timestamps of **last 20 full drive writes**
  - Self-recovery effect diminishes after 20 P/E cycles

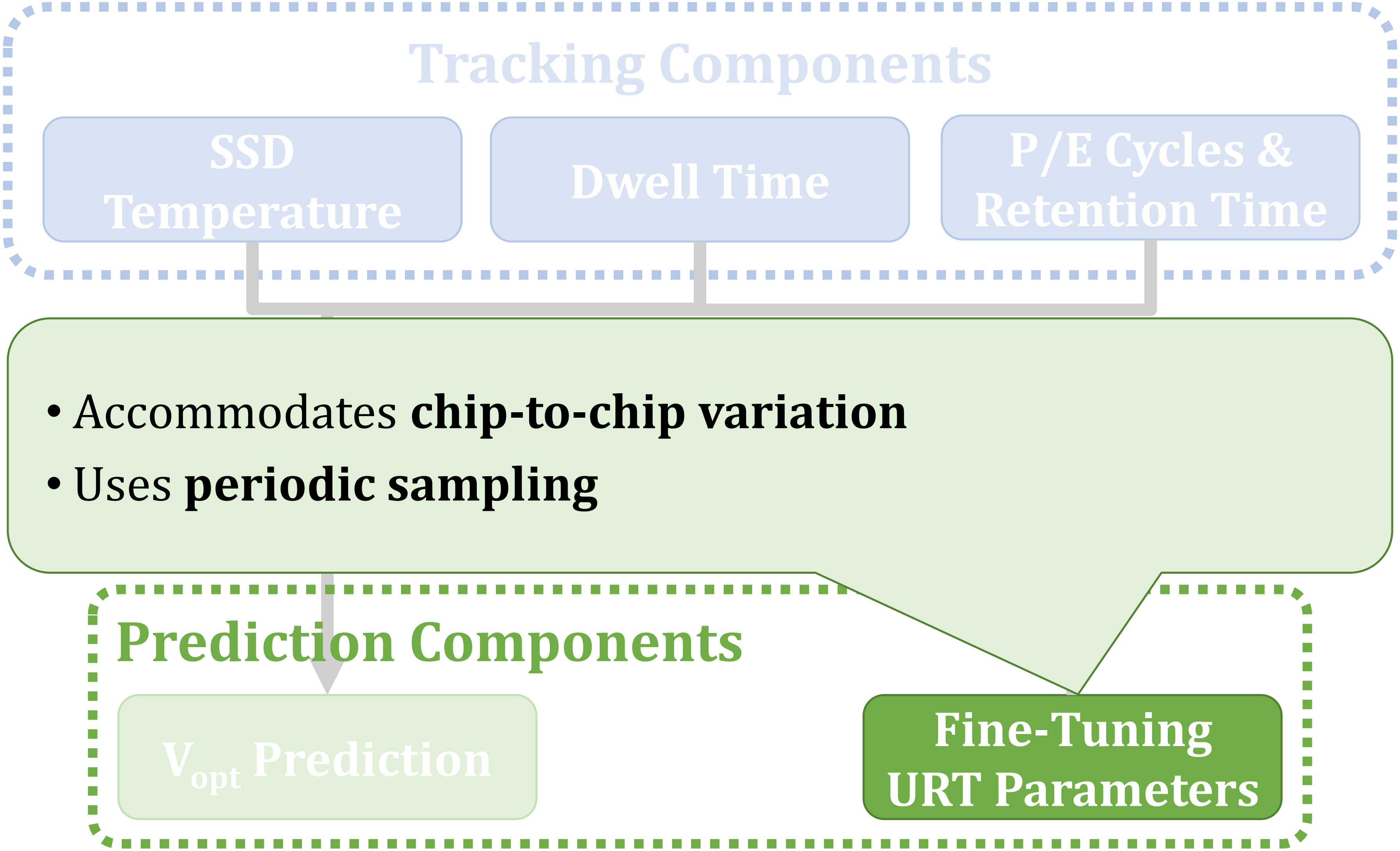
Tracking P/E Cycles and Retention Time



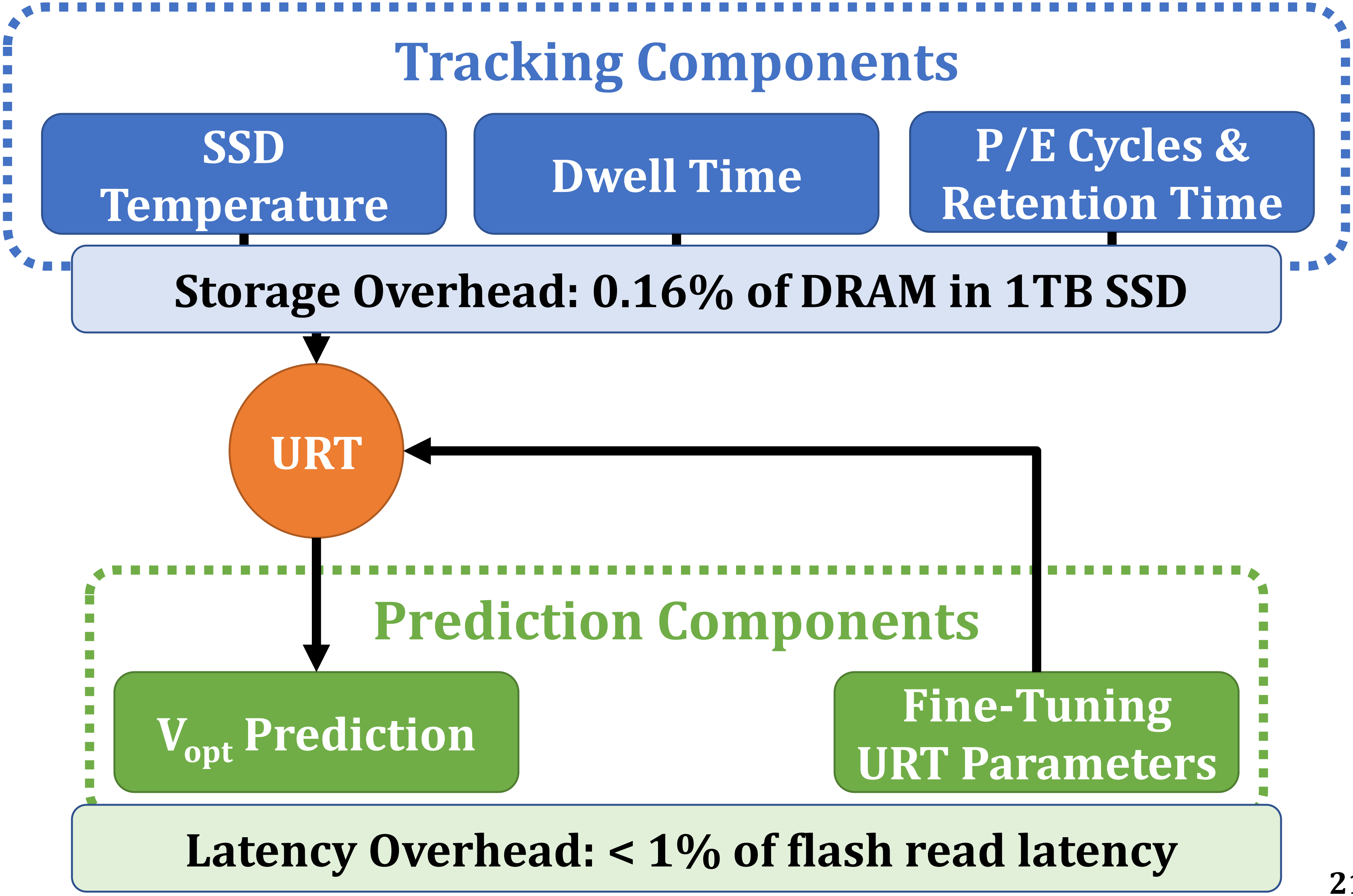
Predicting Optimal Read Reference Voltage



Fine-Tuning URT Parameters Online



HeatWatch Mechanism Summary

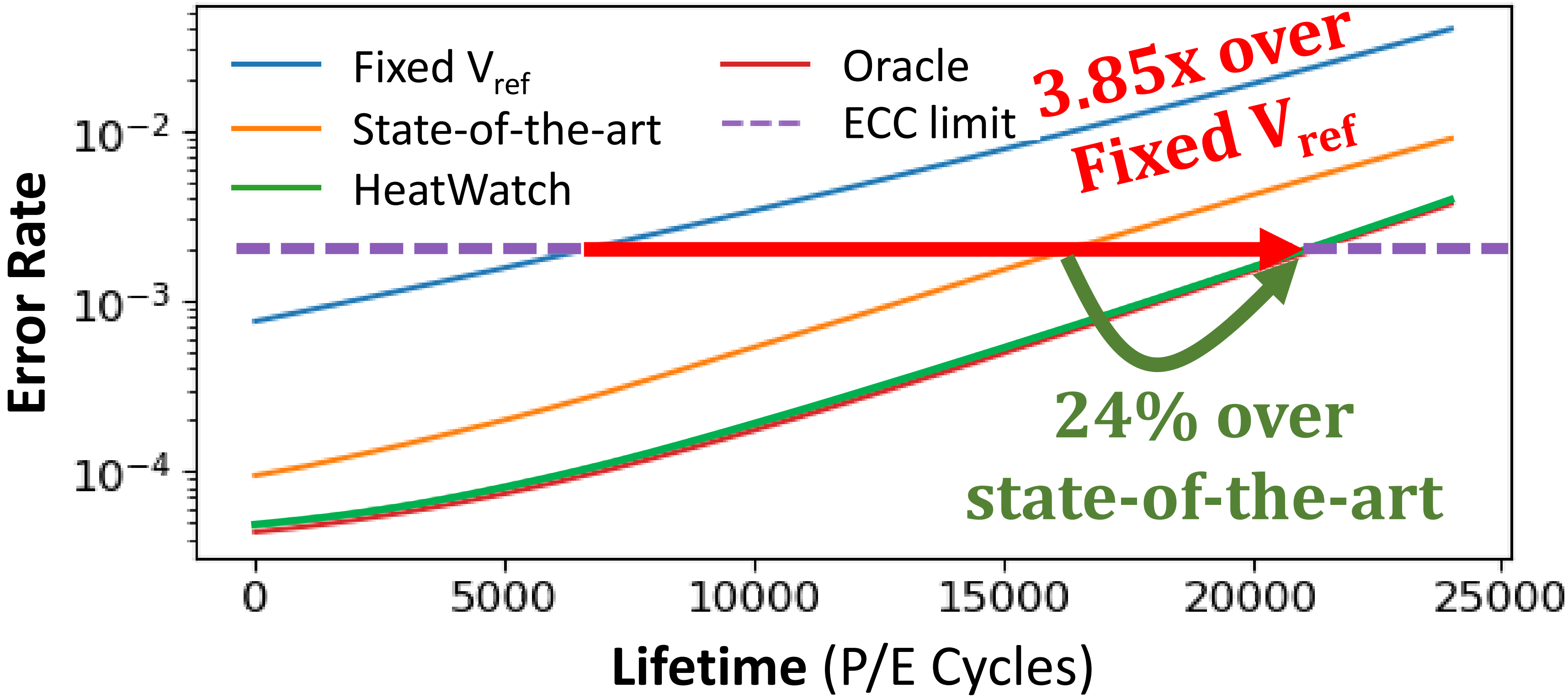


HeatWatch Evaluation Methodology

- **28 real workload storage traces**
  - MSR-Cambridge
  - We use **real dwell time, retention time values** obtained from traces

- **Temperature Model:**
  Trigonometric function + Gaussian noise
  - Represents periodic temperature variation in each day
  - Includes small transient temperature variation

HeatWatch Greatly Improves Flash Lifetime



HeatWatch improves lifetime by capturing the effect of retention, wearout, self-recovery, temperature

Outline

- Executive Summary
- Background on NAND Flash Reliability
- Characterization of Self-Recovery and Temperature Effect on Real 3D NAND Flash Memory Chips
- URT: Unified Self-Recovery and Temperature Model
- HeatWatch Mechanism
- **Conclusion**

Conclusion

- 3D NAND flash memory susceptible to **retention errors**
  - Charge leaks out of flash cell
  - Two unreported factors: ***self-recovery* and *temperature***
- We study *self-recovery* and *temperature* effects
  - **Experimental characterization** of *real* 3D NAND chips
  - **Unified Self-Recovery and Temperature (URT) Model**
    - Predicts impact of retention loss, wearout, self-recovery, temperature on **flash cell voltage**
    - **Low prediction error rate: 4.9%**
- We develop a new technique to improve flash reliability
  - **HeatWatch**
    - Uses URT model to find optimal read voltages for 3D NAND flash
    - **Improves flash lifetime by 3.85x**

HeatWatch

Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature Awareness

Yixin Luo **Saugata Ghose** **Yu Cai** **Erich F. Haratsch** **Onur Mutlu**

Carnegie Mellon
SK hynix
ETH Zürich
SAFARI
SEAGATE