Computer Architecture

Lecture 5: Memory Security and Reliability (in 2020)

Prof. Onur Mutlu

ETH Zürich

Fall 2020

1 October 2020

Four Key Problems + Directions

- Fundamentally Secure/Reliable/Safe Architectures

- Fundamentally Energy-Efficient Architectures
  - Memory-centric (Data-centric) Architectures

- Fundamentally Low-Latency and Predictable Architectures

- Architectures for AI/ML, Genomics, Medicine, Health

Security Implications

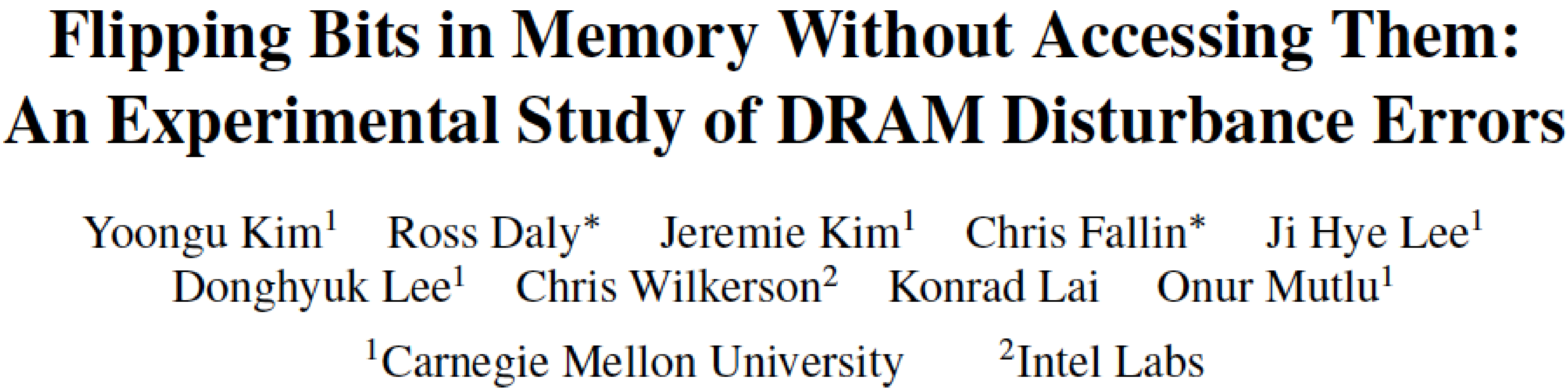
It's like breaking into an apartment by repeatedly slamming a neighbor's door until the vibrations open the door you were after

Understanding RowHammer

RowHammer Solutions

First RowHammer Analysis

- Yoongu Kim, Ross Daly, Jeremie Kim, Chris Fallin, Ji Hye Lee, Donghyuk Lee, Chris Wilkerson, Konrad Lai, and Onur Mutlu, **"Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors"** *Proceedings of the 41st International Symposium on Computer Architecture* (***ISCA***), Minneapolis, MN, June 2014. [Slides (pptx) (pdf)] [Lightning Session Slides (pptx) (pdf)] [Source Code and Data]



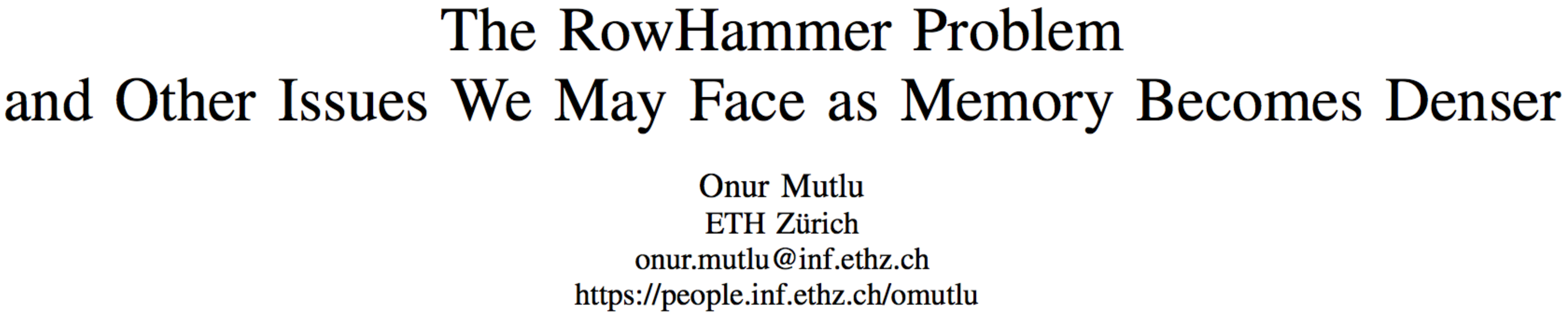
Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors

Yoongu Kim[1] Ross Daly* Jeremie Kim[1] Chris Fallin* Ji Hye Lee[1]
Donghyuk Lee[1] Chris Wilkerson[2] Konrad Lai Onur Mutlu[1]

[1]Carnegie Mellon University [2]Intel Labs

Retrospective on RowHammer & Future

- Onur Mutlu,
  "The RowHammer Problem and Other Issues We May Face as Memory Becomes Denser"
  *Invited Paper in Proceedings of the Design, Automation, and Test in Europe Conference (**DATE**)*, Lausanne, Switzerland, March 2017.
  [Slides (pptx) (pdf)]



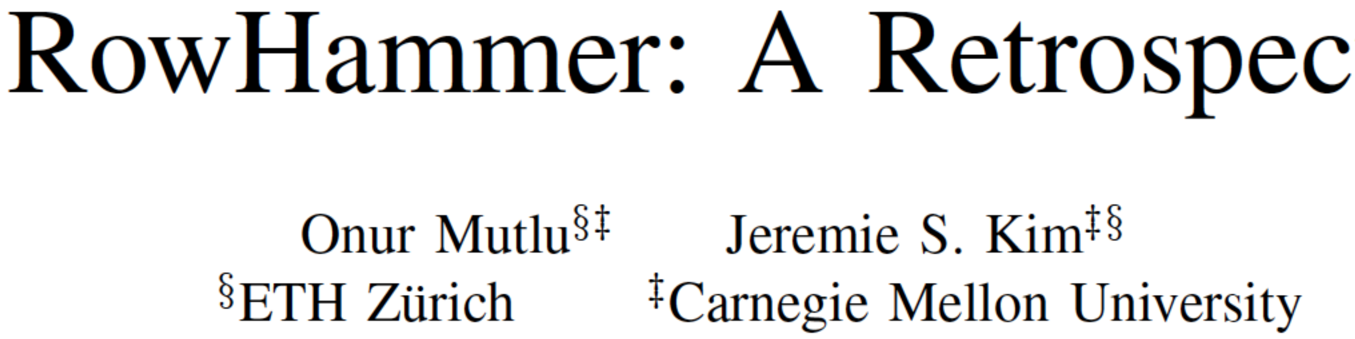
The RowHammer Problem
and Other Issues We May Face as Memory Becomes Denser

Onur Mutlu
ETH Zürich
onur.mutlu@inf.ethz.ch
https://people.inf.ethz.ch/omutlu

https://people.inf.ethz.ch/omutlu/pub/rowhammer-and-other-memory-issues_date17.pdf

A More Recent RowHammer Retrospective

- Onur Mutlu and Jeremie Kim,
  "RowHammer: A Retrospective"
  *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**) Special Issue on Top Picks in Hardware and Embedded Security*, 2019.
  [Preliminary arXiv version]



RowHammer: A Retrospective

Onur Mutlu[§‡] Jeremie S. Kim[‡§]
[§]ETH Zürich [‡]Carnegie Mellon University

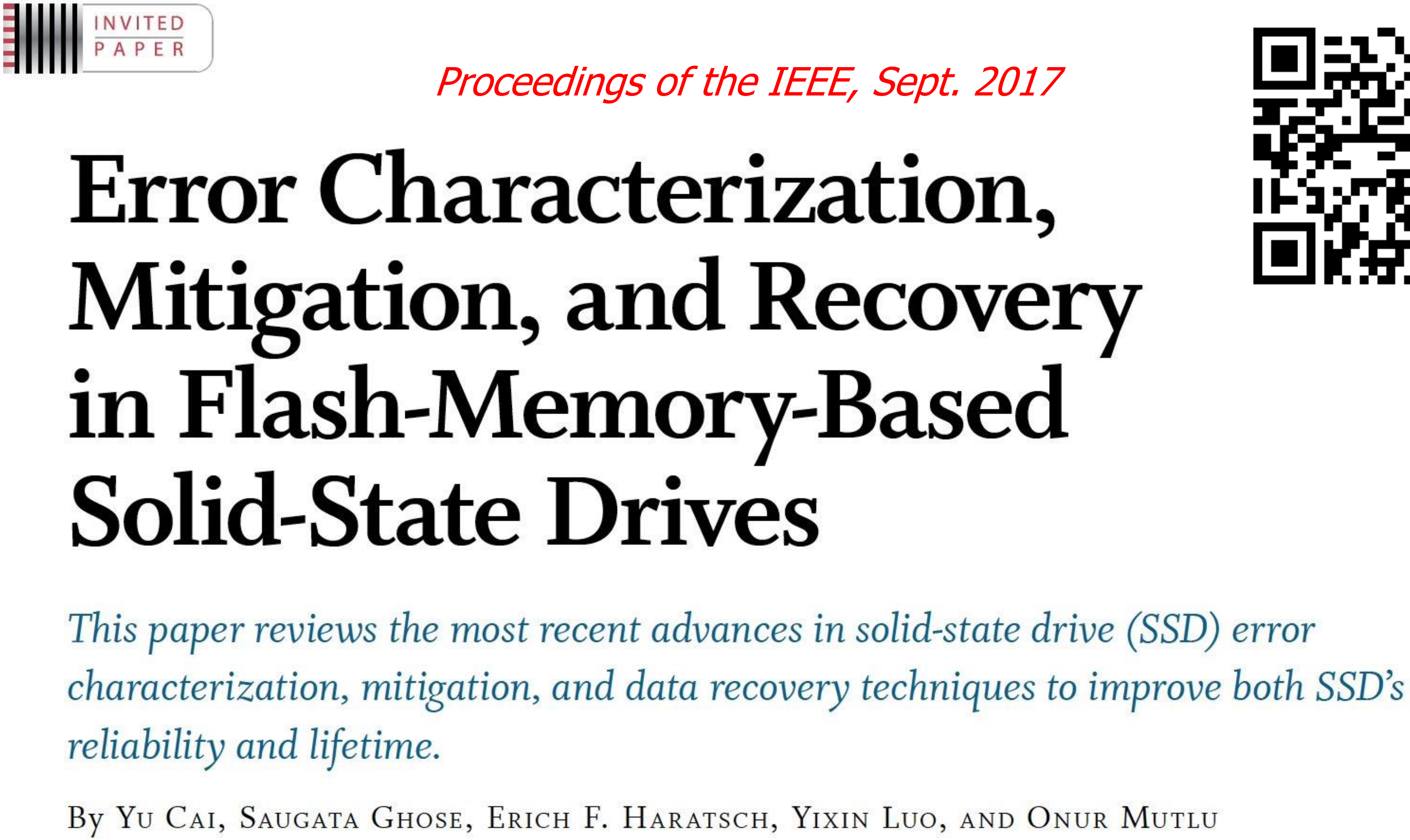
A Key Takeaway

Main Memory Needs

Intelligent Controllers

Aside: Intelligent Controller for NAND Flash

Proceedings of the IEEE, Sept. 2017



INVITED PAPER

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.

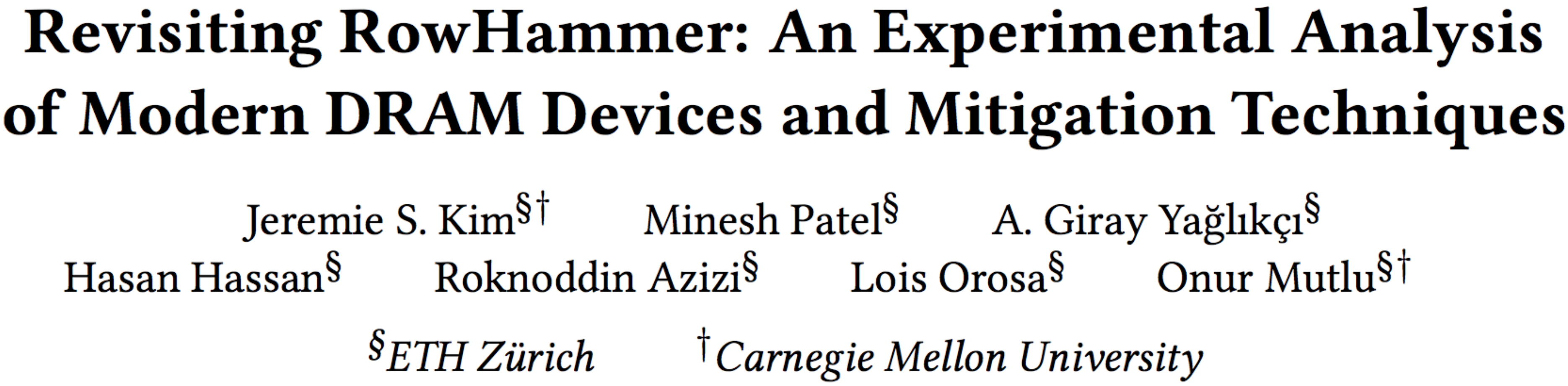
By Yu Cai, Saugata Ghose, Erich F. Haratsch, Yixin Luo, and Onur Mutlu

https://arxiv.org/pdf/1706.08642

RowHammer in 2020

RowHammer in 2020 (I)

- Jeremie S. Kim, Minesh Patel, A. Giray Yaglikci, Hasan Hassan, Roknoddin Azizi, Lois Orosa, and Onur Mutlu,
  "Revisiting RowHammer: An Experimental Analysis of Modern Devices and Mitigation Techniques"
  Proceedings of the 47th International Symposium on Computer Architecture (***ISCA***), Valencia, Spain, June 2020.
  [Slides (pptx) (pdf)]
  [Lightning Talk Slides (pptx) (pdf)]
  [Talk Video (20 minutes)]
  [Lightning Talk Video (3 minutes)]



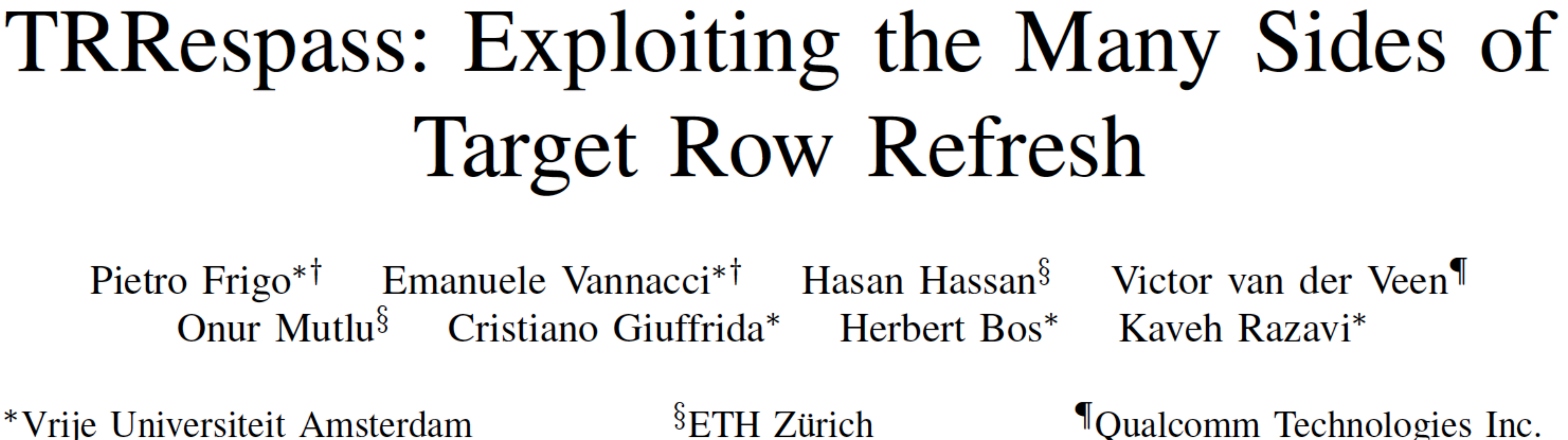
Revisiting RowHammer: An Experimental Analysis of Modern DRAM Devices and Mitigation Techniques

Jeremie S. Kim[§†] Minesh Patel[§] A. Giray Yağlıkçı[§]
Hasan Hassan[§] Roknoddin Azizi[§] Lois Orosa[§] Onur Mutlu[§†]

[§]*ETH Zürich* [†]*Carnegie Mellon University*

RowHammer in 2020 (II)

- Pietro Frigo, Emanuele Vannacci, Hasan Hassan, Victor van der Veen, Onur Mutlu, Cristiano Giuffrida, Herbert Bos, and Kaveh Razavi, **"TRRespass: Exploiting the Many Sides of Target Row Refresh"** *Proceedings of the 41st IEEE Symposium on Security and Privacy (**S&P**)*, San Francisco, CA, USA, May 2020.
  [Slides (pptx) (pdf)]
  [Talk Video (17 minutes)]
  [Source Code]
  [Web Article]
  Best paper award.



TRRespass: Exploiting the Many Sides of Target Row Refresh

Pietro Frigo[*†] Emanuele Vannacci[*†] Hasan Hassan[§] Victor van der Veen[¶]
Onur Mutlu[§] Cristiano Giuffrida[*] Herbert Bos[*] Kaveh Razavi[*]

[*]Vrije Universiteit Amsterdam [§]ETH Zürich [¶]Qualcomm Technologies Inc.

RowHammer in 2020 (III)

- Lucian Cojocar, Jeremie Kim, Minesh Patel, Lillian Tsai, Stefan Saroiu, Alec Wolman, and Onur Mutlu, **"Are We Susceptible to Rowhammer? An End-to-End Methodology for Cloud Providers"** *Proceedings of the 41st IEEE Symposium on Security and Privacy (**S&P**)*, San Francisco, CA, USA, May 2020.
[Slides (pptx) (pdf)]
[Talk Video (17 minutes)]

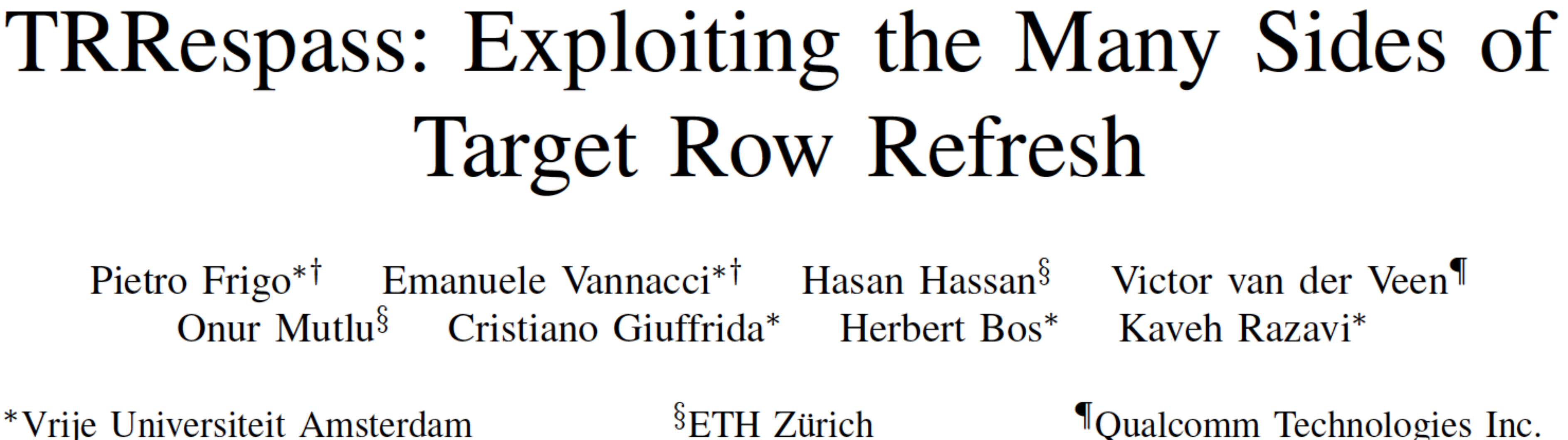
Are We Susceptible to Rowhammer?
An End-to-End Methodology for Cloud Providers

Lucian Cojocar, Jeremie Kim[§†], Minesh Patel[§], Lillian Tsai[‡],
Stefan Saroiu, Alec Wolman, and Onur Mutlu[§†]
Microsoft Research, [§]ETH Zürich, [†]CMU, [‡]MIT

TRRespass

RowHammer in 2020 (II)

- Pietro Frigo, Emanuele Vannacci, Hasan Hassan, Victor van der Veen, Onur Mutlu, Cristiano Giuffrida, Herbert Bos, and Kaveh Razavi, **"TRRespass: Exploiting the Many Sides of Target Row Refresh"** *Proceedings of the 41st IEEE Symposium on Security and Privacy (**S&P**)*, San Francisco, CA, USA, May 2020.
  [Slides (pptx) (pdf)]
  [Talk Video (17 minutes)]
  [Source Code]
  [Web Article]
  Best paper award.



TRRespass: Exploiting the Many Sides of Target Row Refresh

Pietro Frigo[*†] Emanuele Vannacci[*†] Hasan Hassan[§] Victor van der Veen[¶]
Onur Mutlu[§] Cristiano Giuffrida[*] Herbert Bos[*] Kaveh Razavi[*]

[*]Vrije Universiteit Amsterdam [§]ETH Zürich [¶]Qualcomm Technologies Inc.

TRRespass

- First work that shows that TRR-protected DRAM chips are vulnerable to RowHammer in the field
  - Mitigations advertised as secure are not secure

- Introduces the Many-sided RowHammer attack
  - Idea: Hammer many rows to bypass TRR mitigations (e.g., by overflowing proprietary TRR tables that detect aggressor rows)

- (Partially) reverse-engineers the TRR and pTRR mitigation mechanisms implemented in DRAM chips and memory controllers

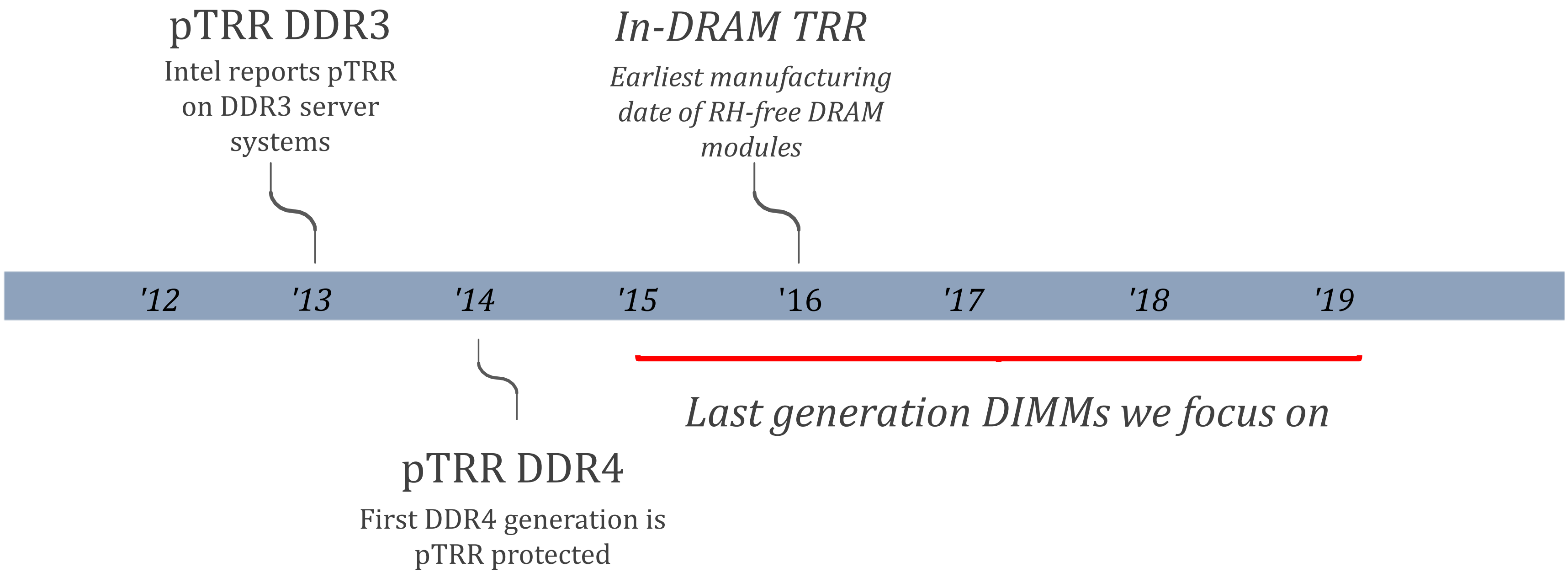
- Provides an automatic tool that can effectively create many-sided RowHammer attacks in DDR4 and LPDDR4(X) chips

Target Row Refresh (TRR)

- How does it work?
  1. *Track activation count of each DRAM row*
  2. *Refresh neighbor rows if row activation count exceeds a threshold*
  - Many possible implementations in practice
  - Security through obscurity

- In-DRAM TRR
  - Embedded in the DRAM circuitry, i.e., not exposed to the memory controller

Timeline of TRR Implementations



pTRR DDR3

Intel reports pTRR on DDR3 server systems

In-DRAM TRR

Earliest manufacturing date of RH-free DRAM modules

Last generation DIMMs we focus on

pTRR DDR4

First DDR4 generation is pTRR protected

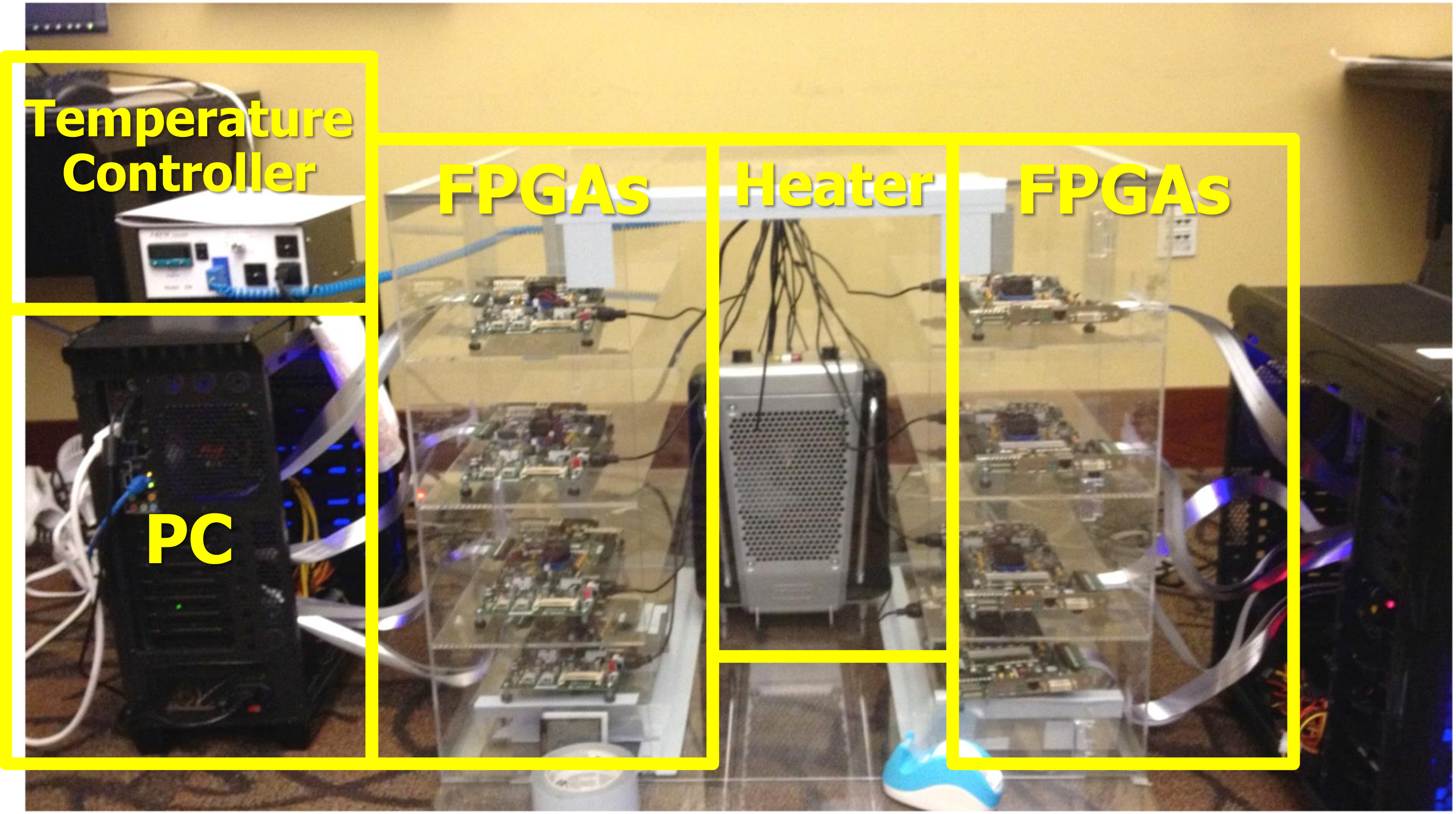
Our Goals

- Reverse engineer in-DRAM TRR to demystify how it works

- Bypass TRR protection
  - A Novel hammering pattern: **The Many-sided RowHammer**
  - Hammering up to **20 aggressor rows** allows bypassing TRR

- Automatically test memory devices: **TRRespass**
  - Automate hammering pattern generation

Infrastructures to Understand Such Issues



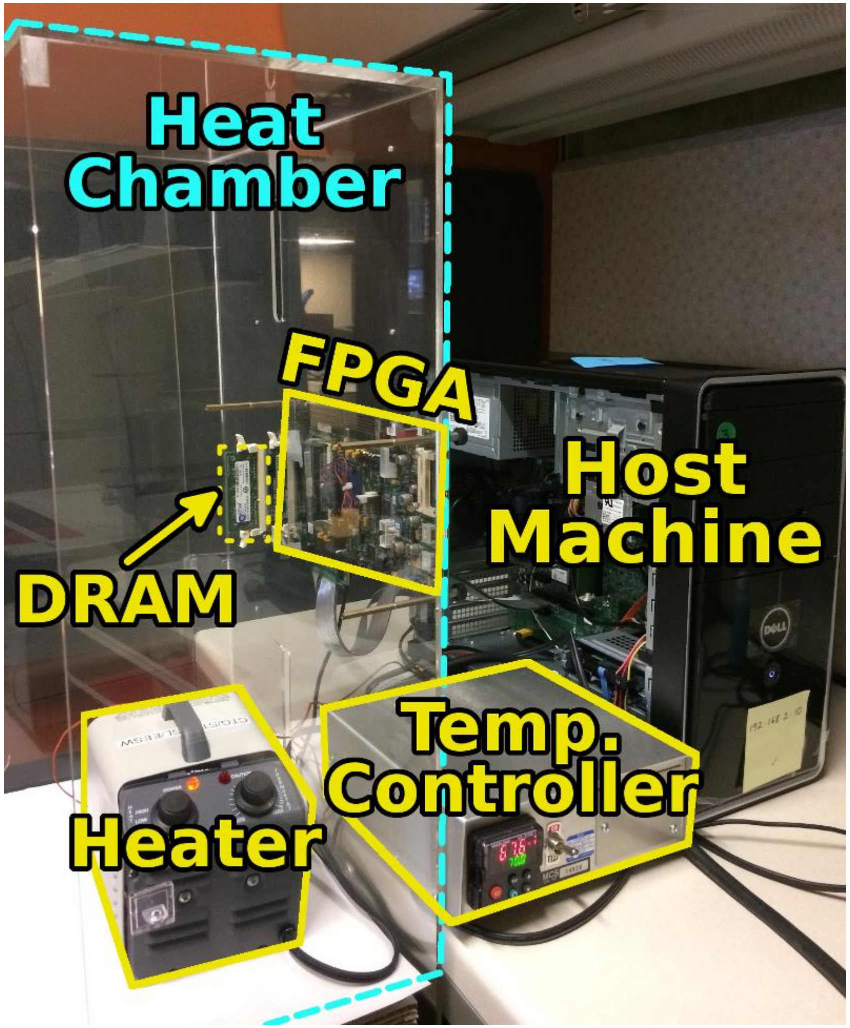
Kim+, "Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors," ISCA 2014.

SoftMC: Open Source DRAM Infrastructure

- Hasan Hassan et al., "**SoftMC: A Flexible and Practical Open-Source Infrastructure for Enabling Experimental DRAM Studies**," HPCA 2017.

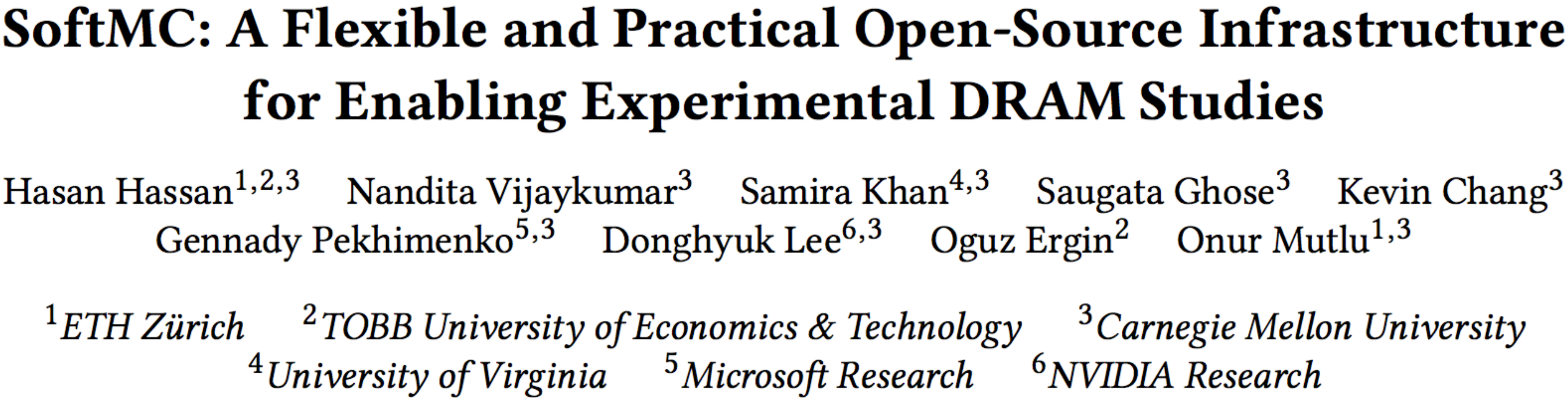
- **Flexible**
- **Easy to Use (C++ API)**
- **Open-source**

  github.com/CMU-SAFARI/SoftMC



SoftMC

- https://github.com/CMU-SAFARI/SoftMC



SoftMC: A Flexible and Practical Open-Source Infrastructure for Enabling Experimental DRAM Studies

Hasan Hassan[1,2,3] Nandita Vijaykumar[3] Samira Khan[4,3] Saugata Ghose[3] Kevin Chang[3]
Gennady Pekhimenko[5,3] Donghyuk Lee[6,3] Oguz Ergin[2] Onur Mutlu[1,3]

[1] *ETH Zürich* [2] *TOBB University of Economics & Technology* [3] *Carnegie Mellon University*
[4] *University of Virginia* [5] *Microsoft Research* [6] *NVIDIA Research*

Components of In-DRAM TRR

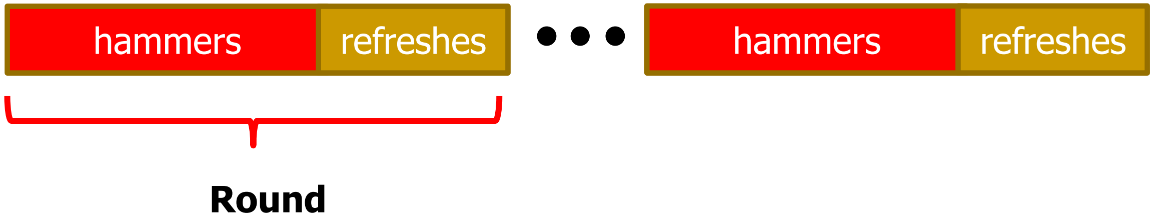
- **Sampler**
  - Tracks aggressor rows activations
  - Design options:
    - Frequency based (record every N^th^ row activation)
    - Time based (record first N row activations)
    - Random seed (record based on a coin flip)
  - Regardless, the sampler has a limited size

- **Inhibitor**
  - Prevents bit flips by refreshing victim rows
    - The latency of performing victim row refreshes is squeezed into slack time available in *tRFC* (i.e., the latency of regular Refresh command)

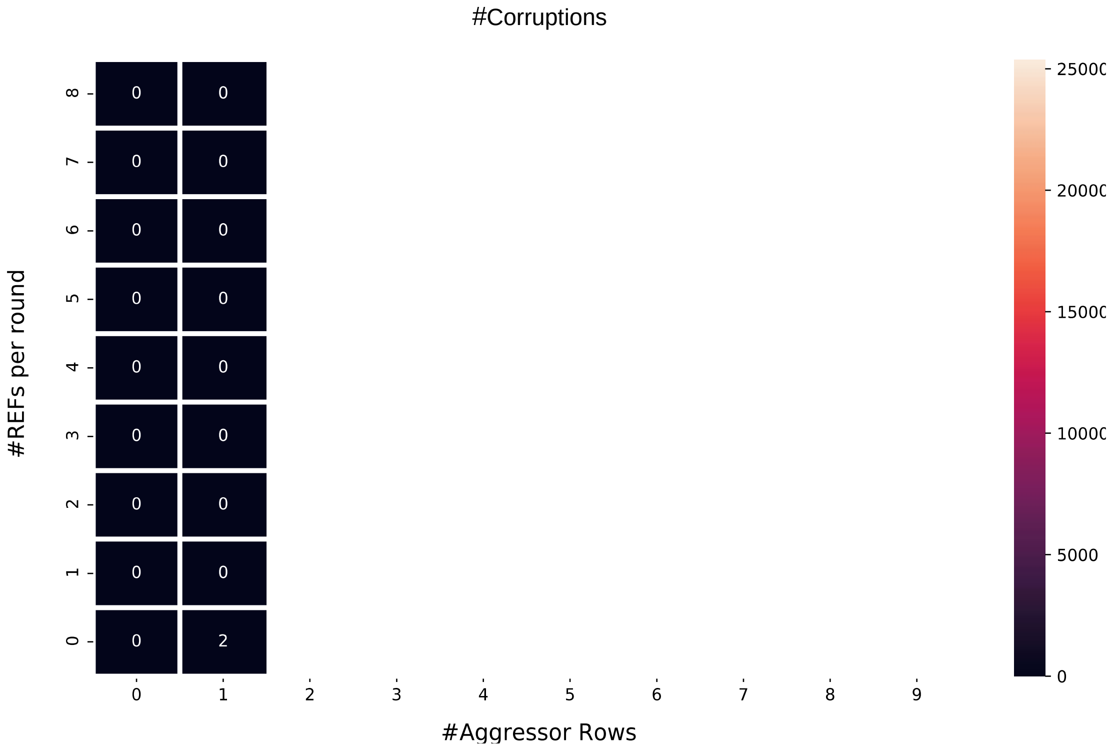
Case Study: Vendor C

How big is the sampler?

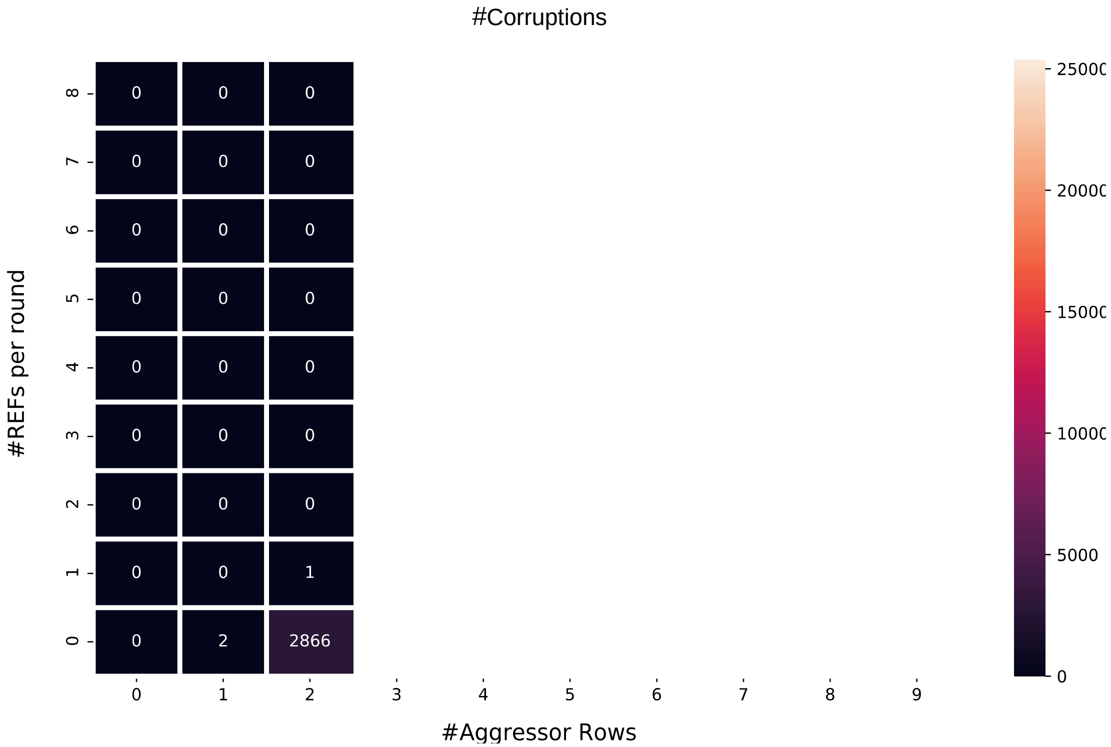
- Pick **N** aggressor rows
- Perform a series of hammers (i.e., activations of aggressors)
  - **8K activations**
- After each series of hammers, issue **R refreshes**
- **10 Rounds**



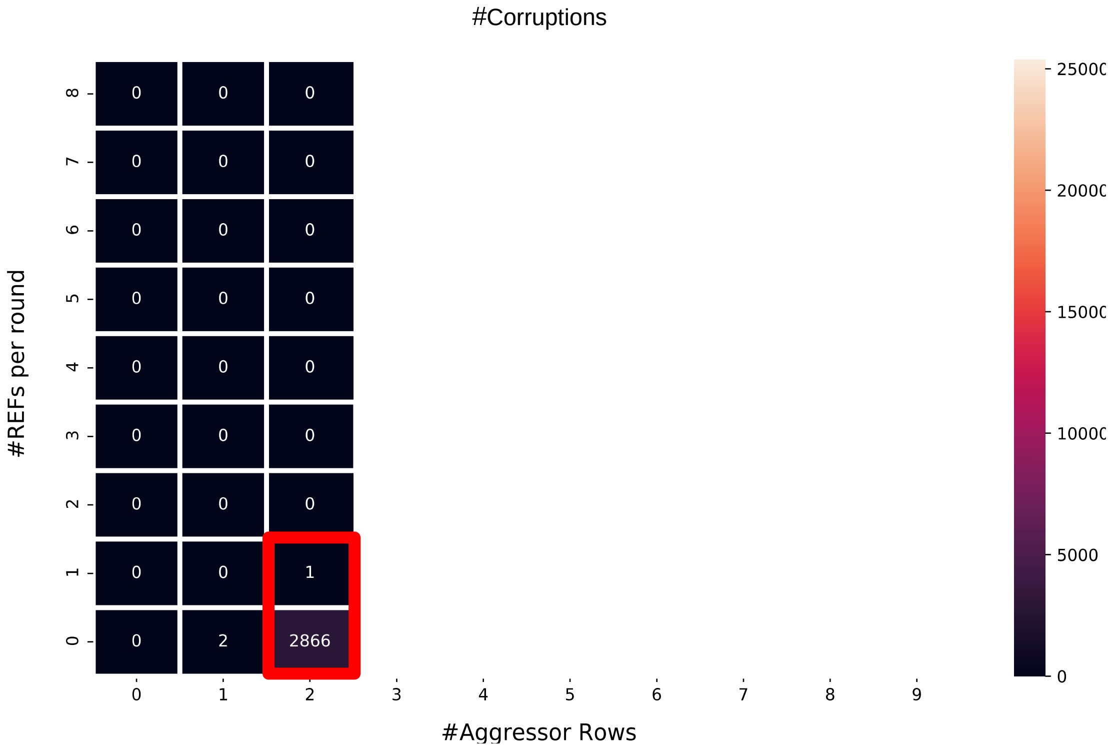
Case Study: Vendor C



Case Study: Vendor C

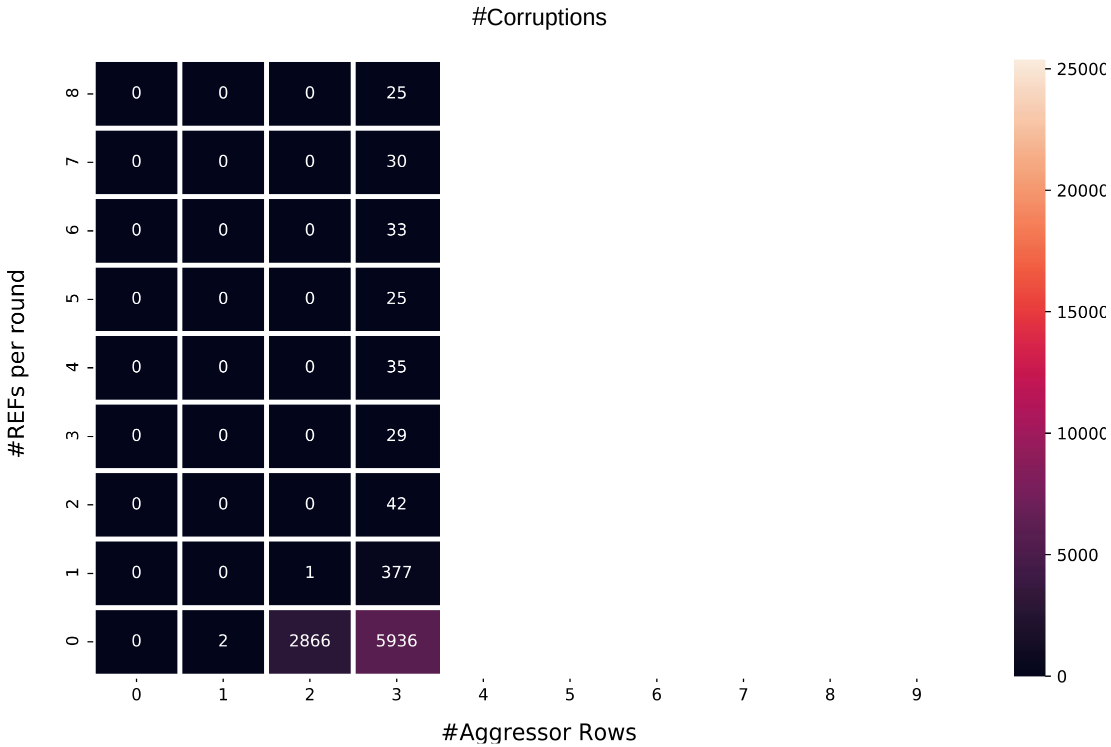


Case Study: Vendor C



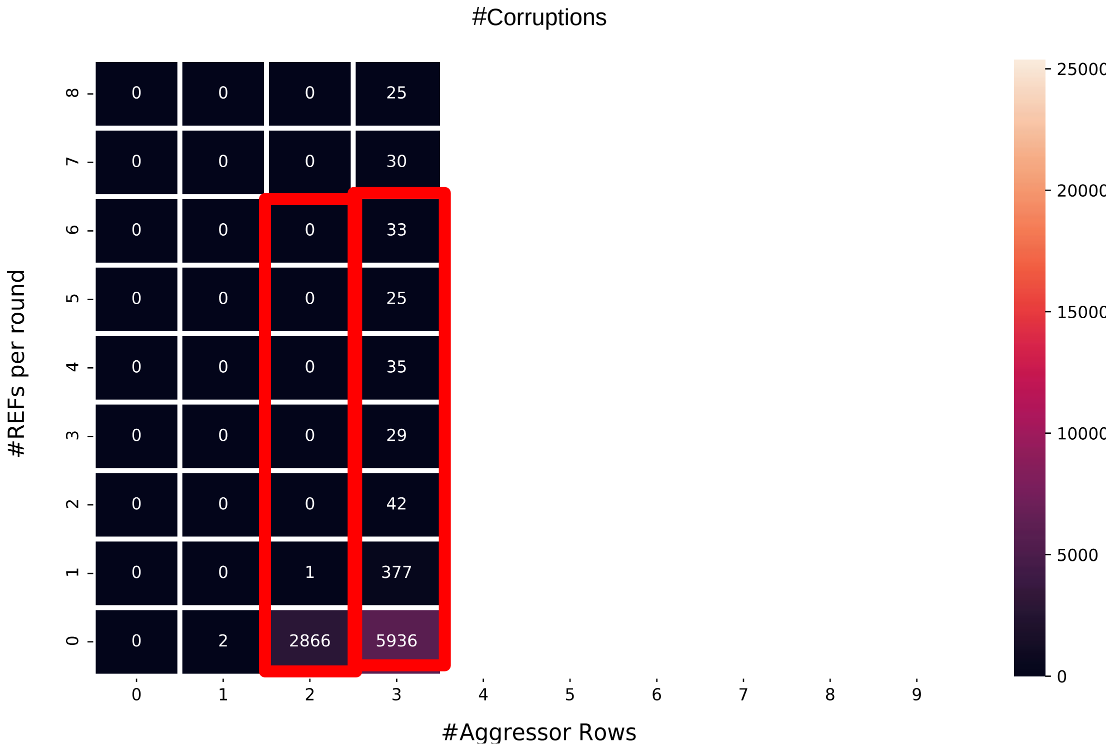
1. The TRR mitigation **acts on a refresh command**

Case Study: Vendor C



#REFs per round	0	1	2	3
8	0	0	0	25
7	0	0	0	30
6	0	0	0	33
5	0	0	0	25
4	0	0	0	35
3	0	0	0	29
2	0	0	0	42
1	0	0	1	377
0	0	2	2866	5936

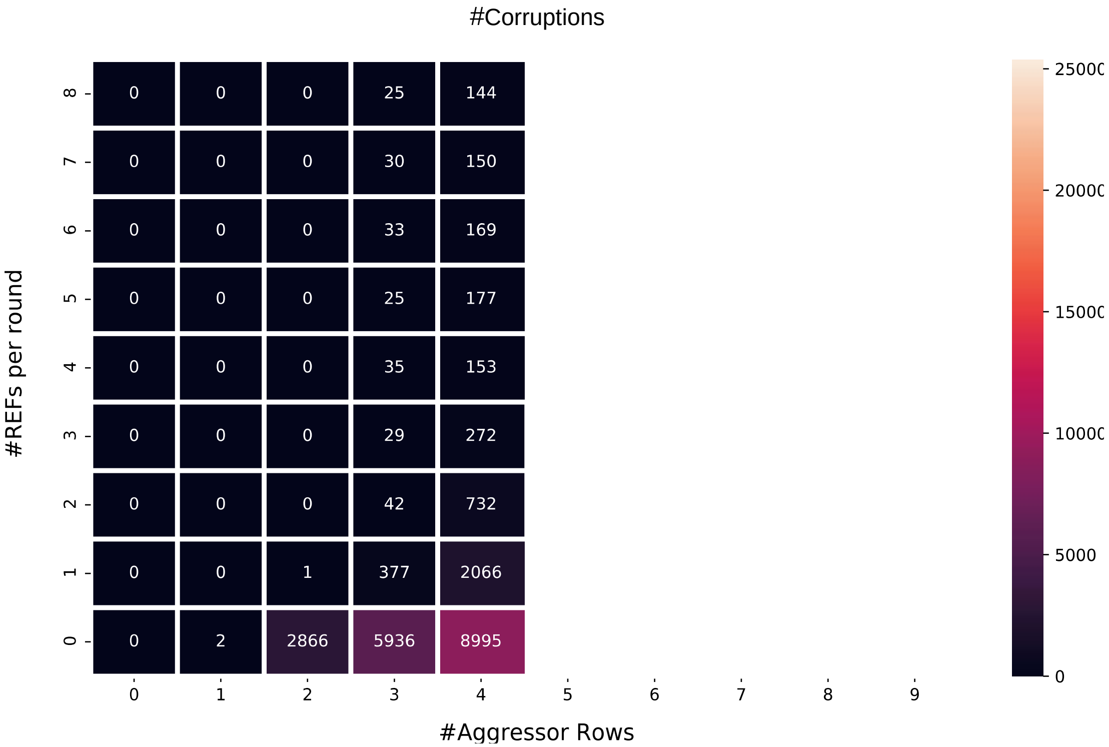
Case Study: Vendor C



2. The mitigation **can sample more than one aggressor** per refresh interval

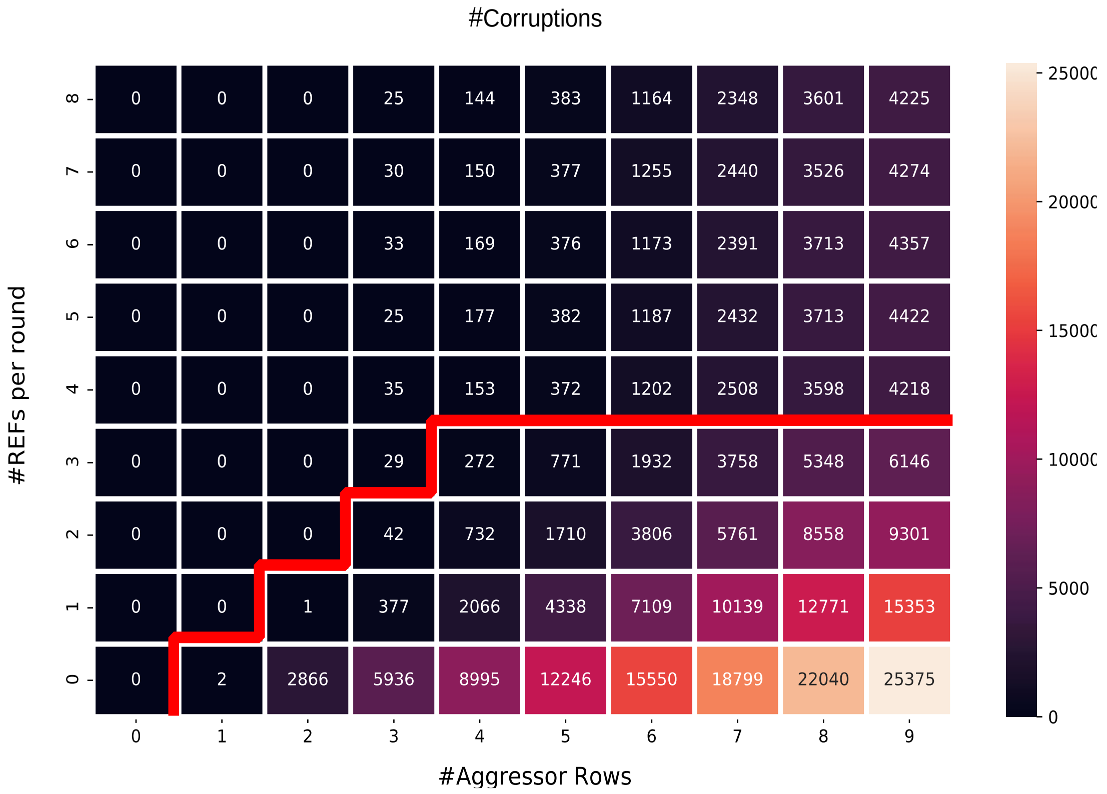
3. The mitigation **can refresh only a single victim** within a refresh operation

Case Study: Vendor C



#REFs per round	0	1	2	3	4	5	6	7	8	9
8	0	0	0	25	144					
7	0	0	0	30	150					
6	0	0	0	33	169					
5	0	0	0	25	177					
4	0	0	0	35	153					
3	0	0	0	29	272					
2	0	0	0	42	732					
1	0	0	1	377	2066					
0	0	2	2866	5936	8995					

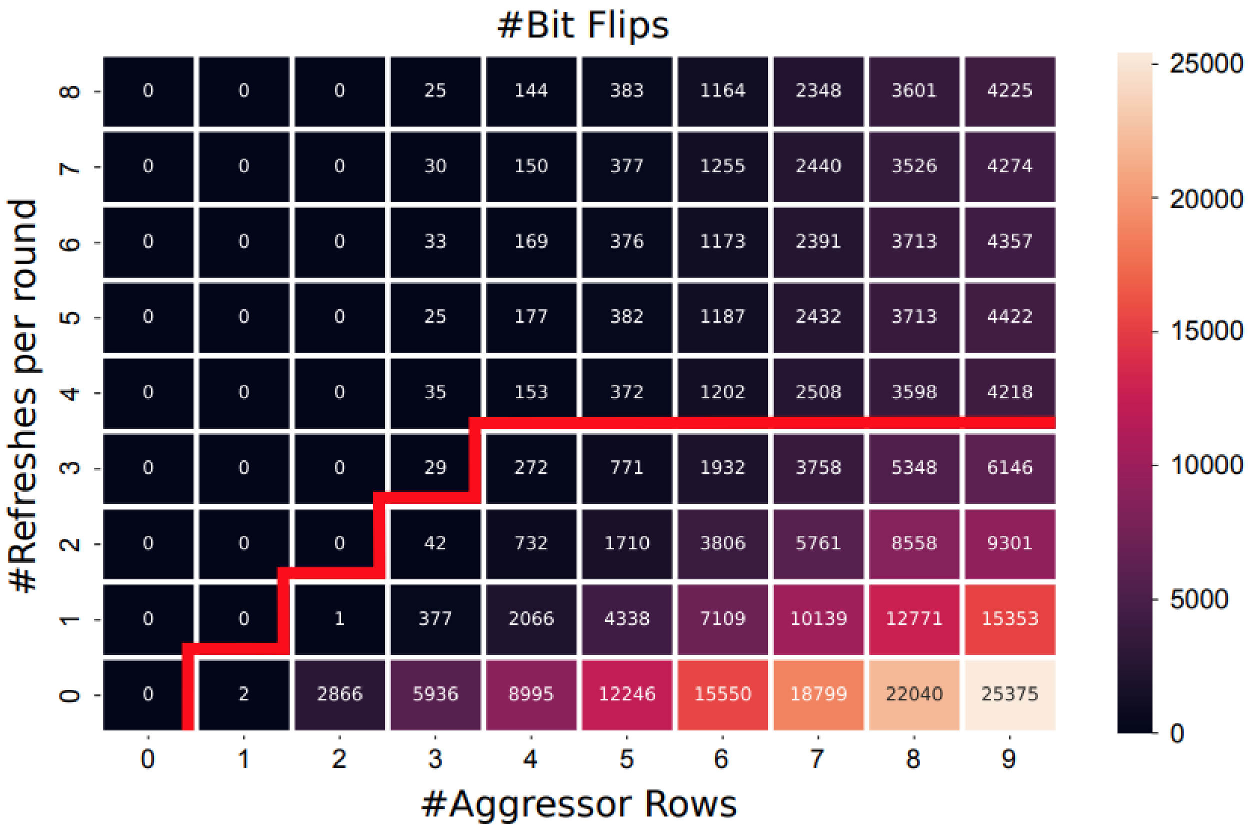
Case Study: Vendor C



#REFs per round \ #Aggressor Rows	0	1	2	3	4	5	6	7	8	9
8	0	0	0	25	144	383	1164	2348	3601	4225
7	0	0	0	30	150	377	1255	2440	3526	4274
6	0	0	0	33	169	376	1173	2391	3713	4357
5	0	0	0	25	177	382	1187	2432	3713	4422
4	0	0	0	35	153	372	1202	2508	3598	4218
3	0	0	0	29	272	771	1932	3758	5348	6146
2	0	0	0	42	732	1710	3806	5761	8558	9301
1	0	0	1	377	2066	4338	7109	10139	12771	15353
0	0	2	2866	5936	8995	12246	15550	18799	22040	25375

4. Sweeping the number of refresh operations and aggressor rows while hammering reveals the sampler size

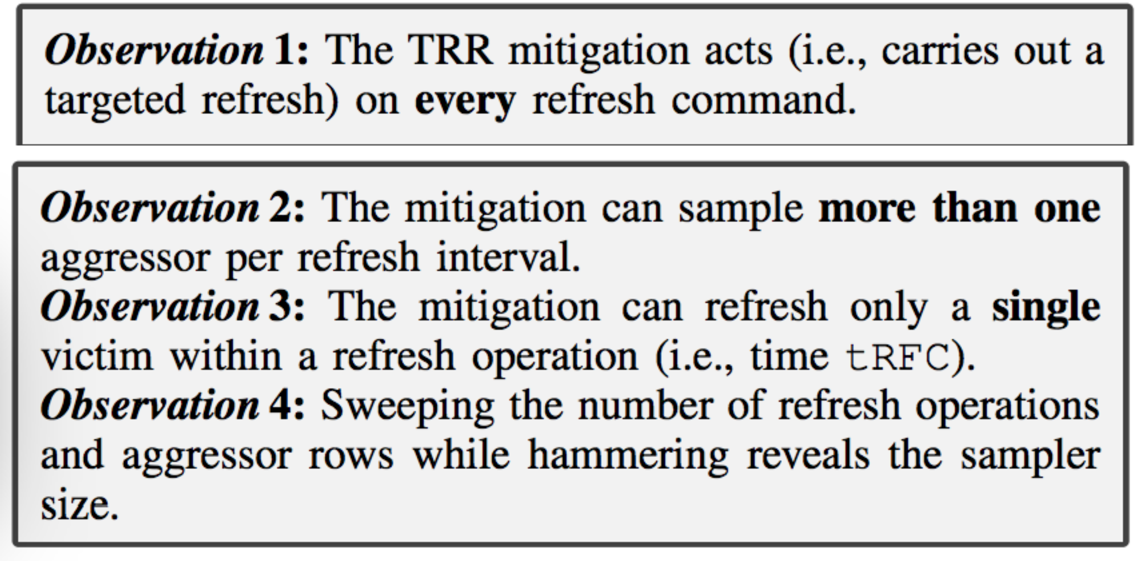
Many-Sided Hammering



	0	1	2	3	4	5	6	7	8	9
8	0	0	0	25	144	383	1164	2348	3601	4225
7	0	0	0	30	150	377	1255	2440	3526	4274
6	0	0	0	33	169	376	1173	2391	3713	4357
5	0	0	0	25	177	382	1187	2432	3713	4422
4	0	0	0	35	153	372	1202	2508	3598	4218
3	0	0	0	29	272	771	1932	3758	5348	6146
2	0	0	0	42	732	1710	3806	5761	8558	9301
1	0	0	1	377	2066	4338	7109	10139	12771	15353
0	0	2	2866	5936	8995	12246	15550	18799	22040	25375

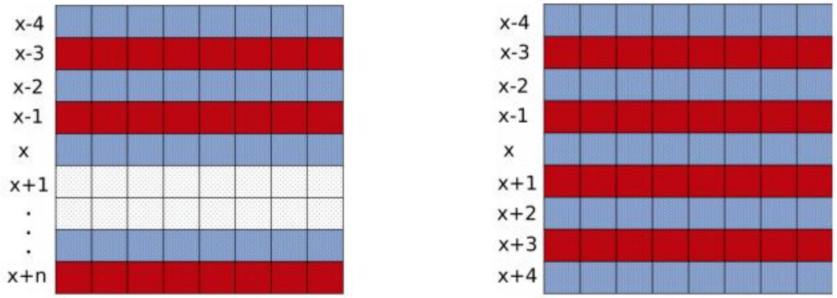
Fig. 9: Refreshes vs. Bit Flips. Module $\mathcal{C}_{12}$: Number of bit flips detected when sending r refresh commands to the module. We report this for different number of aggressor rows (n). For example, when hammering 5 rows, followed by sending 2 refreshes, we find 1,710 bit flips. This figure shows that the number of bit flips stabilizes for $r \geq 4$, implying that the size of the sampler may be 4.

Some Observations



***Observation* 1:** The TRR mitigation acts (i.e., carries out a targeted refresh) on **every** refresh command.

***Observation* 2:** The mitigation can sample **more than one** aggressor per refresh interval.
***Observation* 3:** The mitigation can refresh only a **single** victim within a refresh operation (i.e., time `tRFC`).
***Observation* 4:** Sweeping the number of refresh operations and aggressor rows while hammering reveals the sampler size.

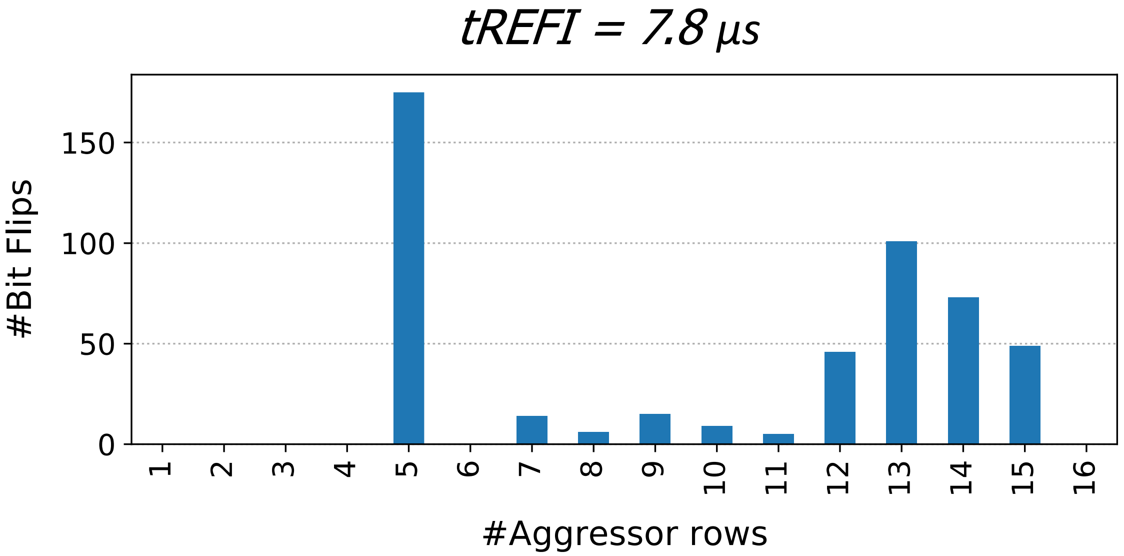


(a) Assisted double-sided **(b)** 4-sided

Fig. 12: Hammering patterns discovered by *TRRespass*. Aggressor rows are in red (🟥) and victim rows are in blue (🟦).

Case Study: Vendor C

Hammering using the default refresh rate



BitFlips vs. Number of Aggressor Rows

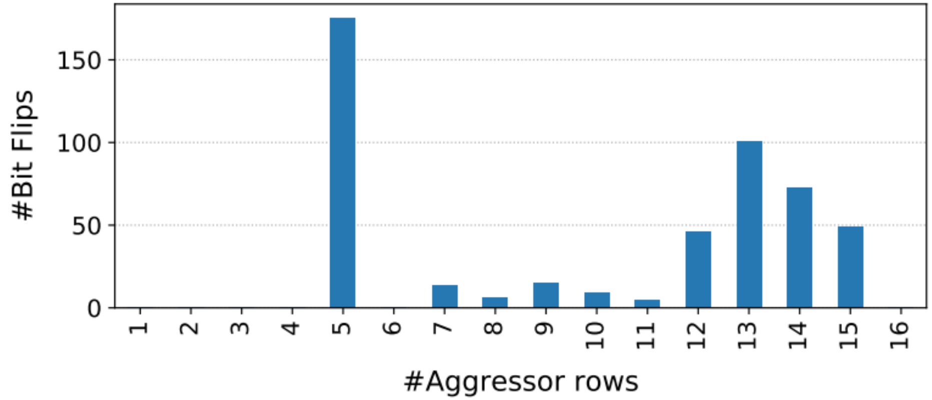


Fig. 10: Bit flips vs. number of aggressor rows. Module $\mathcal{C}_{12}$: Number of bit flips in bank 0 as we vary the number of aggressor rows. Using SoftMC, we refresh DRAM with standard `tREFI` and run the tests until each aggressor rows is hammered 500K times.

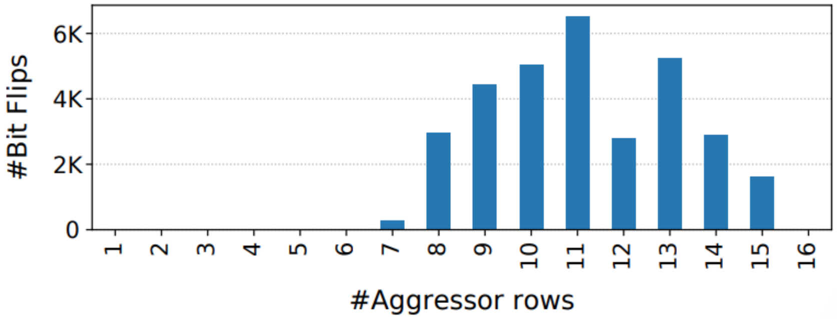


Fig. 11: Bit flips vs. number of aggressor rows. Module $\mathcal{A}_{15}$: Number of bit flips in bank 0 as we vary the number of aggressor rows. Using SoftMC, we refresh DRAM with standard `tREFI` and run the tests until each aggressor rows is hammered 500K times.

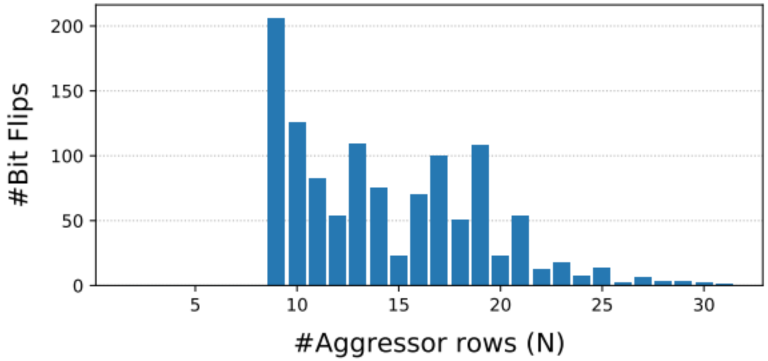


Fig. 13: Bit flips vs. number of aggressor rows. Module $\mathcal{A}_{10}$: Number of bit flips triggered with *N-sided* RowHammer for varying number of *N* on Intel Core i7-7700K. Each aggressor row is one row away from the closest aggressor row (i.e., `VAVAVA...` configuration) and aggressor rows are hammered in a round-robin fashion.

TRRespass Key Results

- 13 out of 42 tested DDR4 DRAM modules are vulnerables
  - From all 3 major manufacturers
  - 3-, 9-, 10-, 14-, 19-sided attacks needed

- 5 out of 13 mobile phones tested vulnerable
  - From 4 major manufacturers
  - With LPDDR4(X) DRAM chips

- These results are scratching the surface
  - TRRespass tool is not exhaustive
  - There is a lot of room for uncovering more vulnerable chips and phones

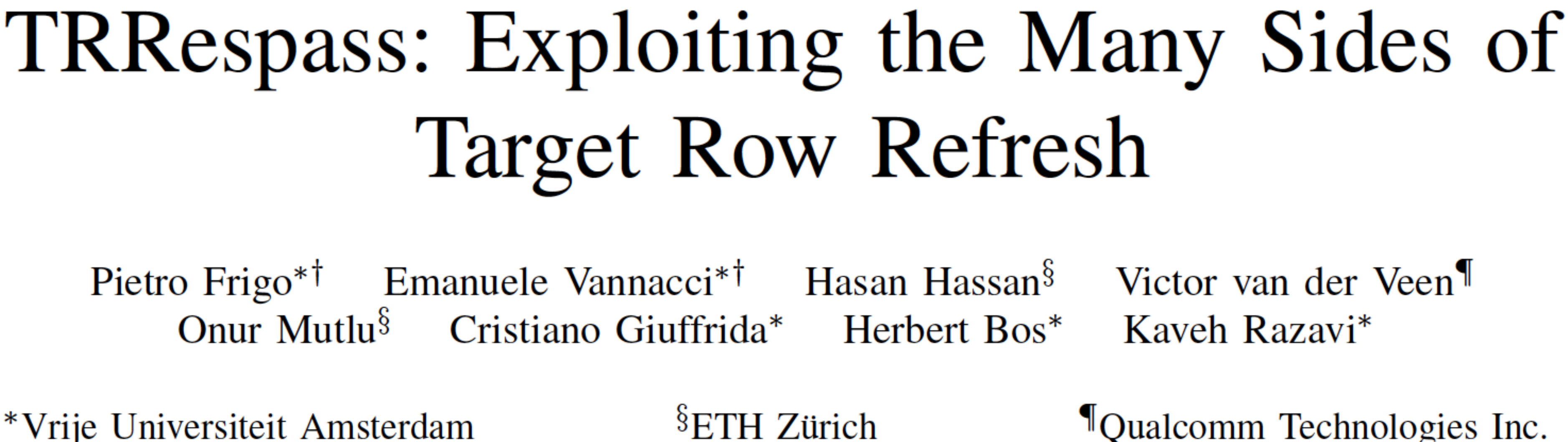
TRRespass Key Takeaways

RowHammer is still an open problem

Security by obscurity is likely not a good solution

More on TRRespass

- Pietro Frigo, Emanuele Vannacci, Hasan Hassan, Victor van der Veen, Onur Mutlu, Cristiano Giuffrida, Herbert Bos, and Kaveh Razavi, **"TRRespass: Exploiting the Many Sides of Target Row Refresh"** *Proceedings of the 41st IEEE Symposium on Security and Privacy (**S&P**)*, San Francisco, CA, USA, May 2020.
  [Slides (pptx) (pdf)]
  [Talk Video (17 minutes)]
  [Source Code]
  [Web Article]
  Best paper award.



TRRespass: Exploiting the Many Sides of Target Row Refresh

Pietro Frigo[*†] Emanuele Vannacci[*†] Hasan Hassan[§] Victor van der Veen[¶]
Onur Mutlu[§] Cristiano Giuffrida[*] Herbert Bos[*] Kaveh Razavi[*]

[*]Vrije Universiteit Amsterdam [§]ETH Zürich [¶]Qualcomm Technologies Inc.

Revisiting RowHammer

RowHammer in 2020 (I)

- Jeremie S. Kim, Minesh Patel, A. Giray Yaglikci, Hasan Hassan, Roknoddin Azizi, Lois Orosa, and Onur Mutlu,
"Revisiting RowHammer: An Experimental Analysis of Modern Devices and Mitigation Techniques"
Proceedings of the 47th International Symposium on Computer Architecture (***ISCA***), Valencia, Spain, June 2020.
[Slides (pptx) (pdf)]
[Lightning Talk Slides (pptx) (pdf)]
[Talk Video (20 minutes)]
[Lightning Talk Video (3 minutes)]

Revisiting RowHammer: An Experimental Analysis of Modern DRAM Devices and Mitigation Techniques

Jeremie S. Kim[§†] Minesh Patel[§] A. Giray Yağlıkçı[§]
Hasan Hassan[§] Roknoddin Azizi[§] Lois Orosa[§] Onur Mutlu[§†]

[§]*ETH Zürich* [†]*Carnegie Mellon University*

Revisiting RowHammer

An Experimental Analysis of Modern Devices and Mitigation Techniques

Jeremie S. Kim Minesh Patel
A. Giray Yağlıkçı Hasan Hassan
Roknoddin Azizi Lois Orosa Onur Mutlu

SAFARI

ETH Zürich

Carnegie Mellon

Executive Summary

- **Motivation**: Denser DRAM chips are more vulnerable to RowHammer but no characterization-based study demonstrates how vulnerability scales
- **Problem**: Unclear if existing mitigation mechanisms will remain viable for future DRAM chips that are likely to be more vulnerable to RowHammer
- **Goal**:
  1. Experimentally demonstrate how vulnerable modern DRAM chips are to RowHammer and study how this vulnerability will scale going forward
  2. Study viability of existing mitigation mechanisms on more vulnerable chips
- **Experimental Study:** First rigorous RowHammer characterization study across a broad range of DRAM chips
  - 1580 chips of different DRAM {types, technology node generations, manufacturers}
  - We find that RowHammer vulnerability worsens in newer chips
- **RowHammer Mitigation Mechanism Study:** How five state-of-the-art mechanisms are affected by worsening RowHammer vulnerability
  - Reasonable performance loss (8% on average) on modern DRAM chips
  - Scale poorly to more vulnerable DRAM chips (e.g., 80% performance loss)
- **Conclusion:** it is critical to research more effective solutions to RowHammer for future DRAM chips that will likely be even more vulnerable to RowHammer

Motivation

- Denser DRAM chips are **more vulnerable** to RowHammer
- Three prior works **[Kim+, ISCA'14], [Park+, MR'16], [Park+, MR'16]**, **over the last six years** provide RowHammer characterization data on real DRAM
- However, there is **no comprehensive experimental study** that demonstrates **how vulnerability scales** across DRAM types and technology node generations
- It is **unclear whether current mitigation mechanisms will remain viable** for future DRAM chips that are likely to be more vulnerable to RowHammer

Goal

1. **Experimentally demonstrate** how vulnerable modern DRAM chips are to RowHammer and **predict how this vulnerability will scale** going forward

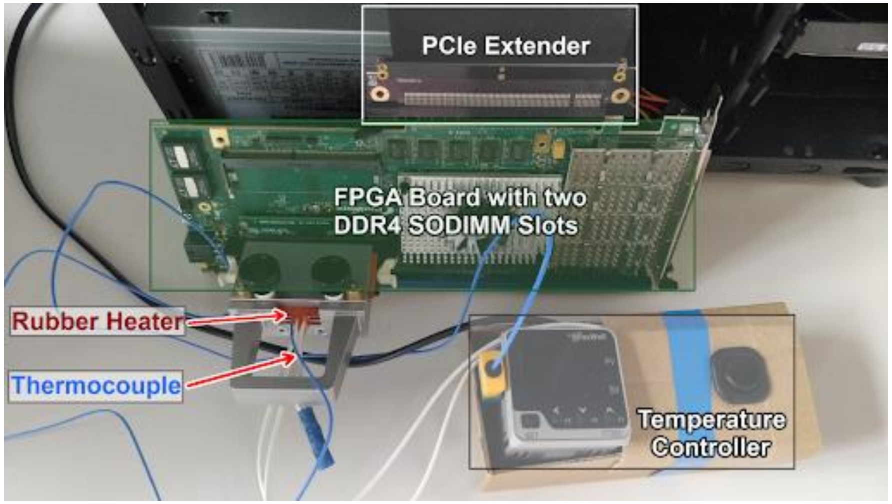
2. Examine the viability of current mitigation mechanisms on **more vulnerable chips**

DRAM Testing Infrastructures

Three separate testing infrastructures

1. **DDR3:** FPGA-based SoftMC [Hassan+, HPCA'17] (Xilinx ML605)
2. **DDR4:** FPGA-based SoftMC [Hassan+, HPCA'17] (Xilinx Virtex UltraScale 95)
3. **LPDDR4:** In-house testing hardware for LPDDR4 chips

All provide fine-grained control over DRAM commands, timing parameters and temperature



DDR4 DRAM testing infrastructure

DRAM Chips Tested

DRAM type-node	Number of Chips (Modules) Tested			
	Mfr. A	*Mfr. B*	*Mfr. C*	*Total*
DDR3-old	56 (10)	88 (11)	28 (7)	**172 (28)**
DDR3-new	80 (10)	52 (9)	104 (13)	**236 (32)**
DDR4-old	112 (16)	24 (3)	128 (18)	**264 (37)**
DDR4-new	264 (43)	16 (2)	108 (28)	**388 (73)**
LPDDR4-1x	12 (3)	180 (45)	N/A	**192 (48)**
LPDDR4-1y	184 (46)	N/A	144 (36)	**328 (82)**

1580 total DRAM chips tested from **300** DRAM modules

- **Three** major DRAM manufacturers {A, B, C}
- **Three** DRAM *types* or *standards* {DDR3, DDR4, LPDDR4}
  - LPDDR4 chips we test implement on-die ECC
- **Two** technology nodes per DRAM type {old/new, 1x/1y}
  - Categorized based on manufacturing date, datasheet publication date, purchase date, and characterization results

Type-node: configuration describing a chip's type and technology node generation: **DDR3-old/new, DDR4-old/new, LPDDR4-1x/1y**

Effective RowHammer Characterization

To characterize our DRAM chips at **worst-case** conditions, we:

1. Prevent sources of interference during core test loop

- We disable:
  - **DRAM refresh**: to avoid refreshing victim row
  - **DRAM calibration events**: to minimize variation in test timing
  - **RowHammer mitigation mechanisms**: to observe circuit-level effects
- Test for **less than refresh window (32ms)** to avoid retention failures

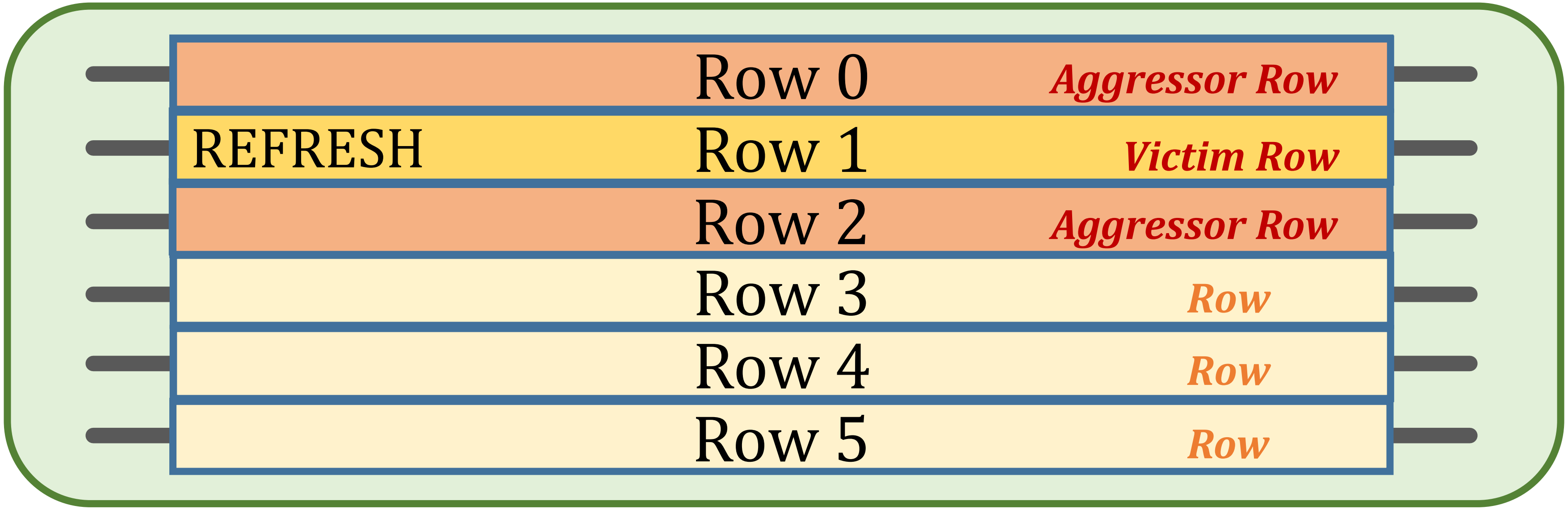
2. Worst-case access sequence

- We use **worst-case** access sequence based on prior works' observations

- For each row, **repeatedly access the two directly physically-adjacent rows as fast as possible**

[More details in the paper]

Testing Methodology

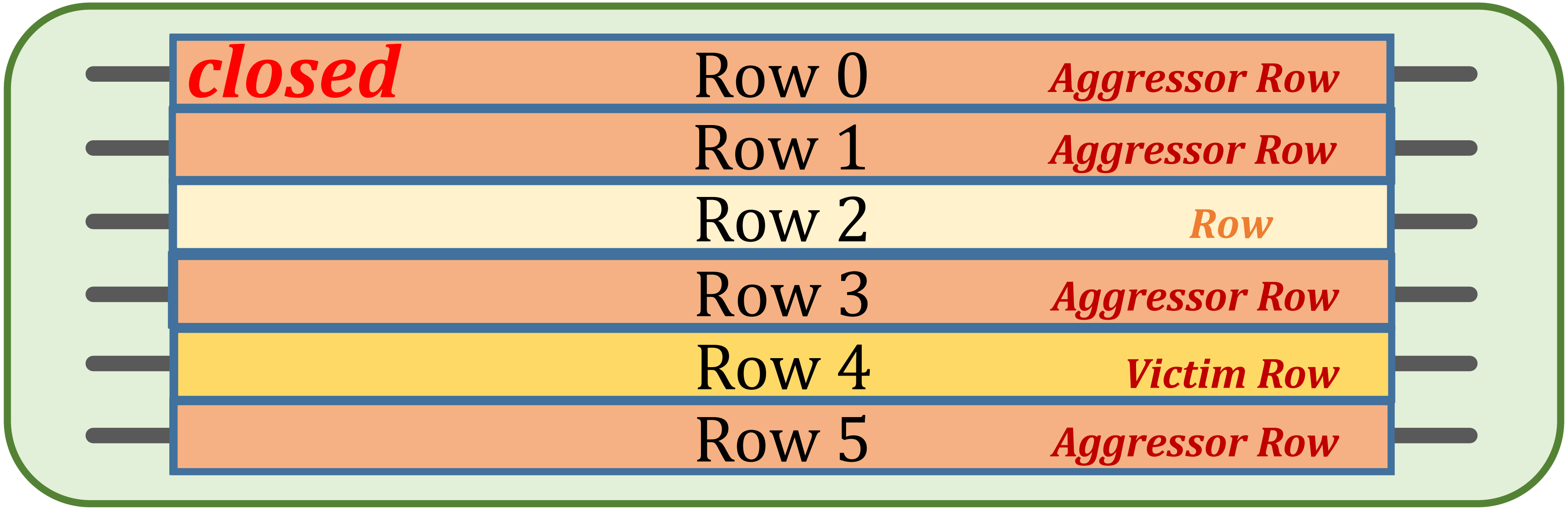

Row 0 Aggressor Row
REFRESH Row 1 Victim Row
Row 2 Aggressor Row
Row 3 Row
Row 4 Row
Row 5 Row

```
DRAM_RowHammer_Characterization():
    foreach row in DRAM:
        set victim_row to row
        set aggressor_row1 to victim_row − 1
        set aggressor_row2 to victim_row + 1
        Disable DRAM refresh
        Refresh victim_row
        for n = 1 → HC: // core test loop
            activate aggressor_row1
            activate aggressor_row2
        Enable DRAM refresh
        Record RowHammer bit flips to storage
        Restore bit flips to original values
```

Disable refresh to **prevent interruptions** in the core loop of our test **from refresh operations**

Induce RowHammer bit flips on a **fully charged row**

Testing Methodology



```
DRAM_RowHammer_Characterization():
    foreach row in DRAM:
        set victim_row to row
        set aggressor_row1 to victim_row – 1
        set aggressor_row2 to victim_row + 1
        Disable DRAM refresh
        Refresh victim_row
        for n = 1 → HC: // core test loop
            activate aggressor_row1
            activate aggressor_row2
        Enable DRAM refresh
        Record RowHammer bit flips to storage
        Restore bit flips to original values
```

Disable refresh to **prevent interruptions** in the core loop of our test **from refresh operations**

Induce RowHammer bit flips on a **fully charged row**

Core test loop where we alternate accesses to adjacent rows

1 Hammer (HC) = two accesses

Prevent further retention failures

Record bit flips for analysis

Key Takeaways from 1580 Chips

- Chips of newer DRAM technology nodes are **more vulnerable** to RowHammer

- There are chips today whose weakest cells fail after **only 4800 hammers**

- Chips of newer DRAM technology nodes can exhibit RowHammer bit flips 1) in **more rows** and 2) **farther away** from the victim row.

1. RowHammer Vulnerability

Q. Can we induce RowHammer bit flips in all of our DRAM chips?

All chips are vulnerable, except many DDR3 chips

- A total of 1320 out of all 1580 chips **(84%)** are vulnerable
- Within **DDR3-old** chips, **only 12%** of chips (24/204) are vulnerable
- Within **DDR3-new** chips, **65%** of chips (148/228) are vulnerable

Newer DRAM chips are more vulnerable to RowHammer

2. Data Pattern Dependence

Q. Are some data patterns more effective in inducing RowHammer bit flips?

- We test **several data patterns** typically examined in prior work to identify the worst-case data pattern

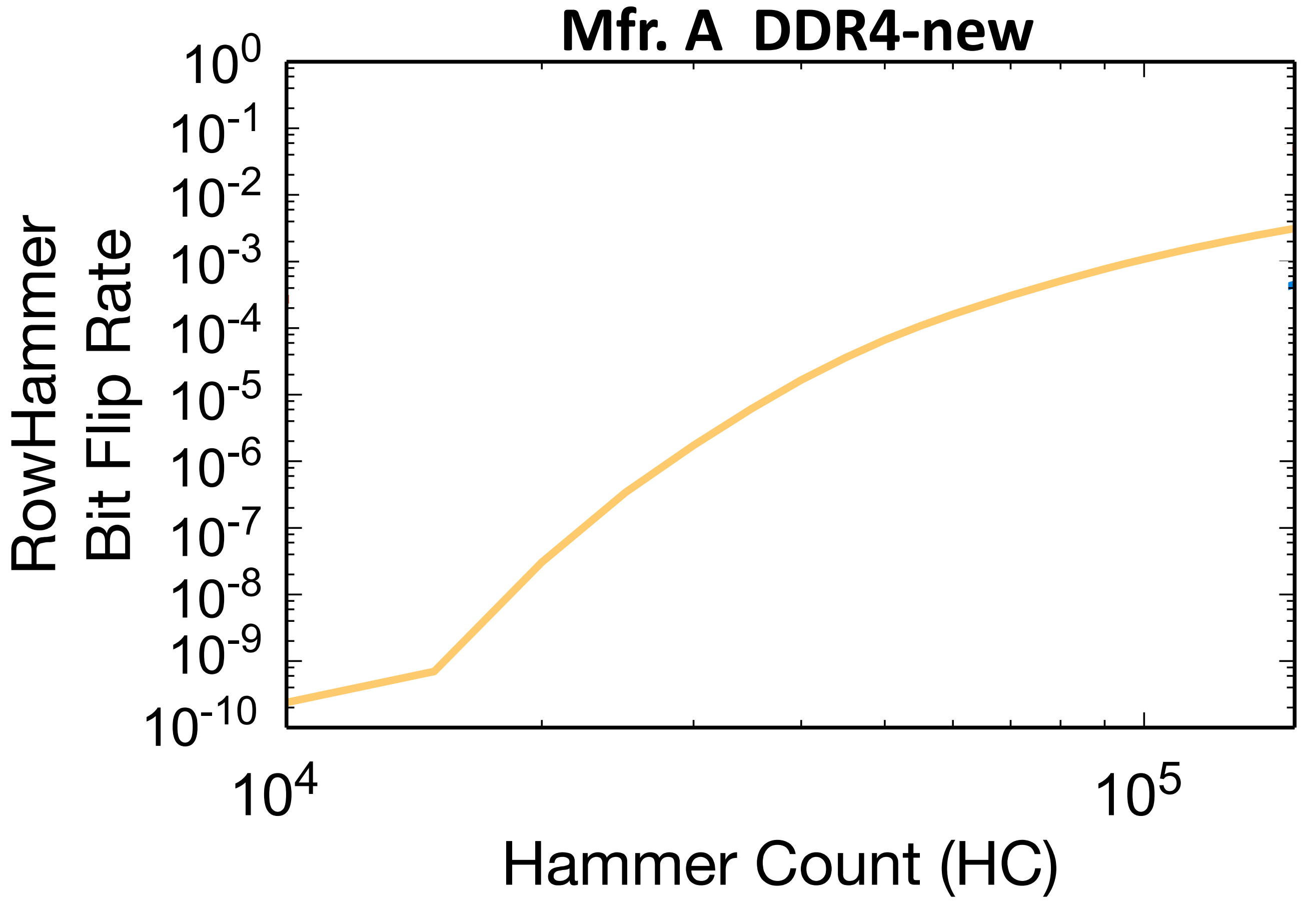
- The worst-case data pattern is **consistent across chips** of the same manufacturer and DRAM type-node configuration

- We use the **worst-case data pattern** per DRAM chip to characterize each chip at **worst-case conditions** and **minimize the extensive testing time**

[More detail and figures in paper]

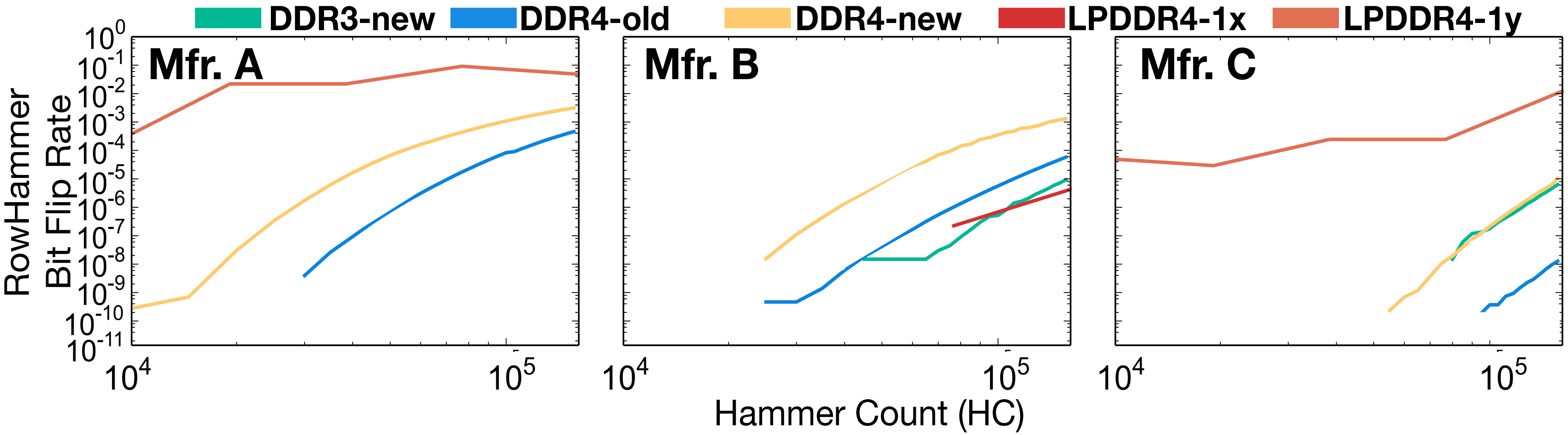
3. Hammer Count (HC) Effects

Q. How does the Hammer Count affect the number of bit flips induced?



Hammer Count = 2 Accesses, one to each adjacent row of victim

3. Hammer Count (HC) Effects

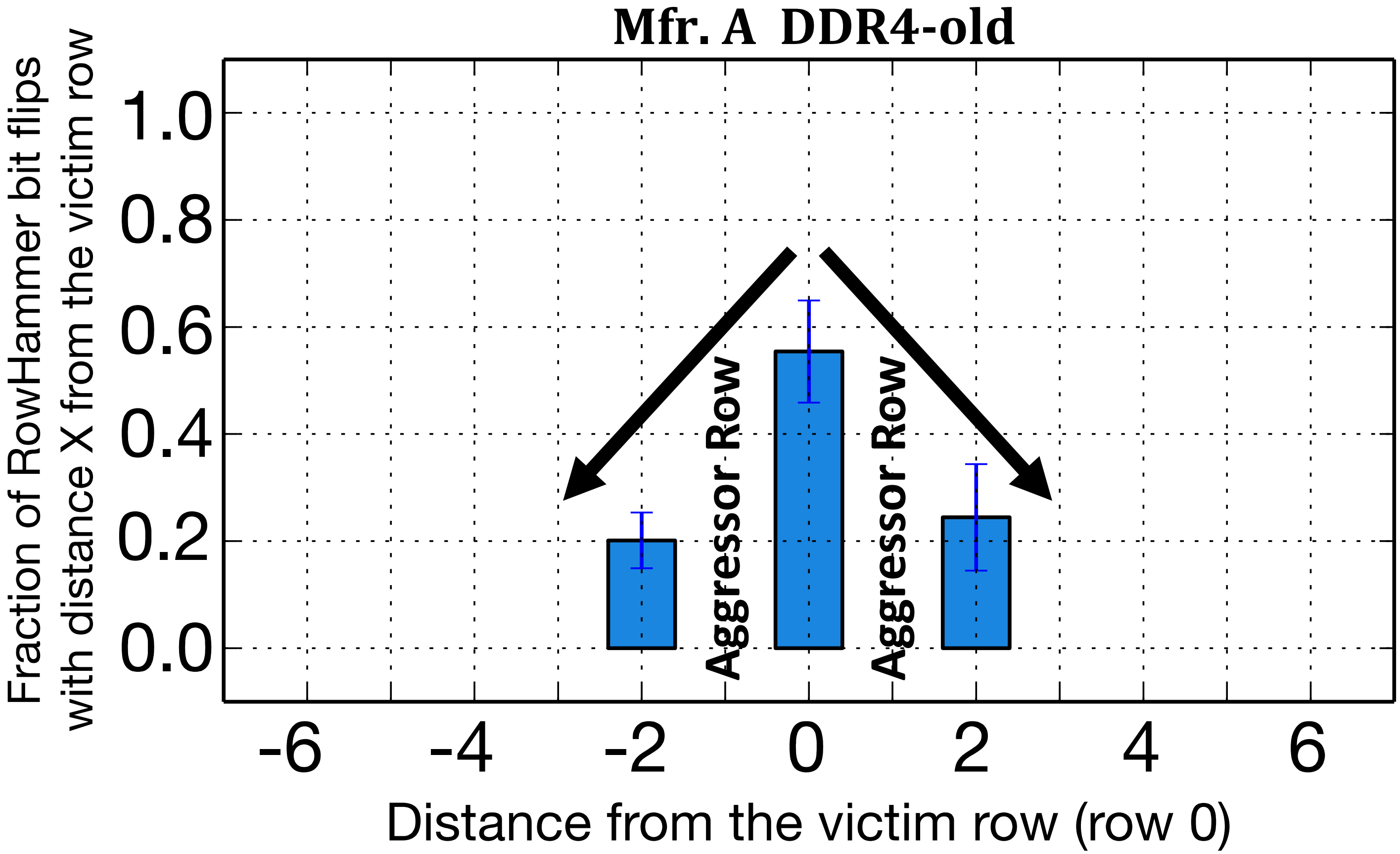


RowHammer bit flip rates **increase**
when going **from old to new** DDR4 technology node generations

RowHammer bit flip rates (i.e., RowHammer vulnerability) increase with technology node generation

4. Spatial Effects: Row Distance

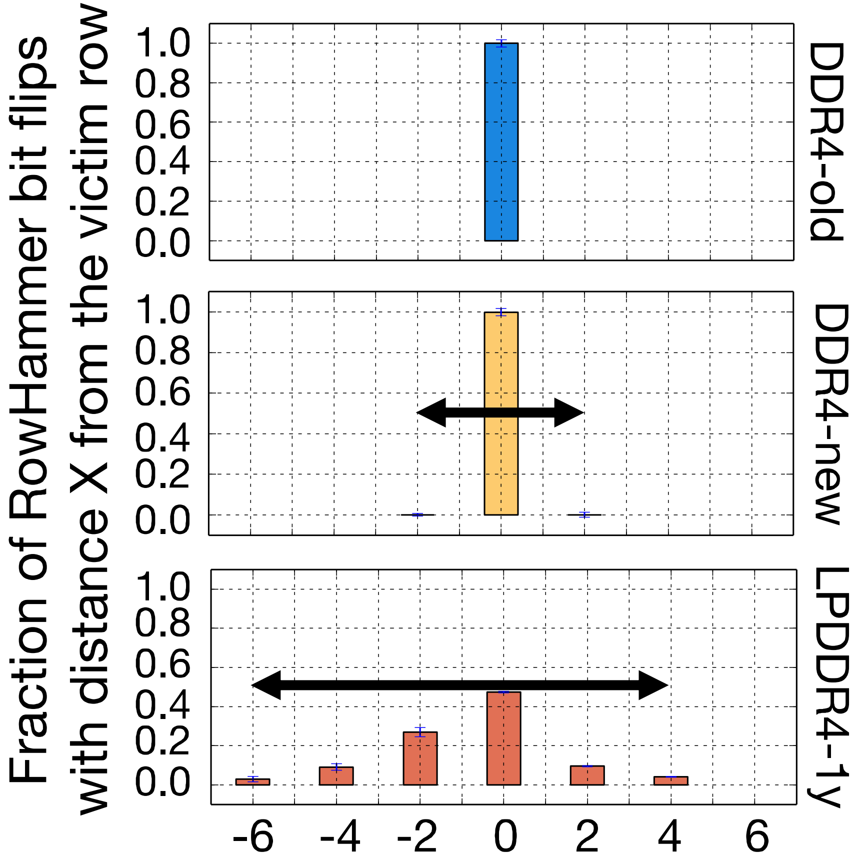
Q. Where do RowHammer bit flips occur relative to aggressor rows?



The number of RowHammer bit flips that occur in a given row decreases as the distance from the **victim row (row 0)** increases.

4. Spatial Effects: Row Distance

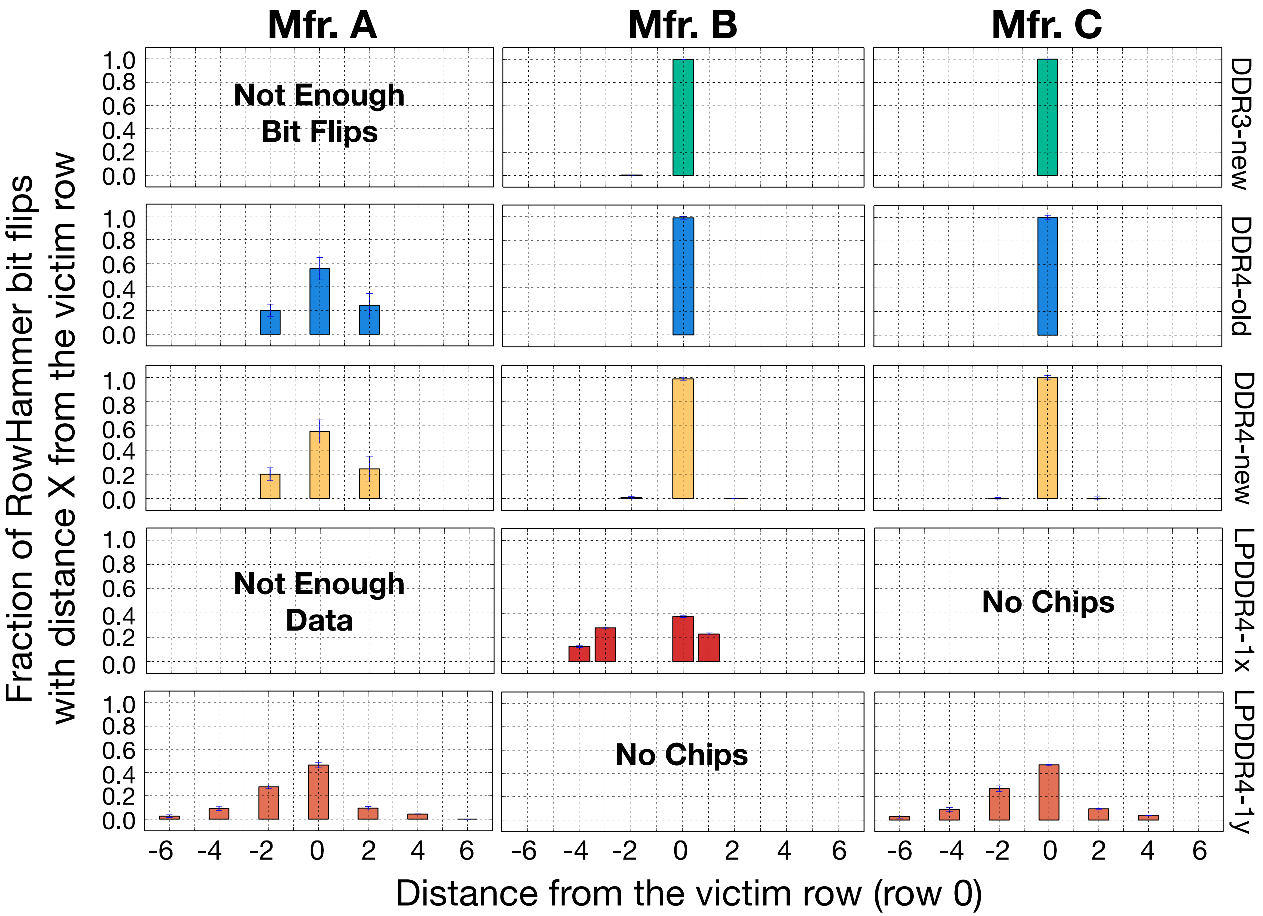
We normalize data by inducing a bit flip rate of 10^{-6} in each chip



Chips of newer DRAM technology nodes can exhibit RowHammer bit flips 1) in **more rows** and 2) **farther away** from the victim row.

4. Spatial Effects: Row Distance

We plot this data for each DRAM type-node configuration per manufacturer



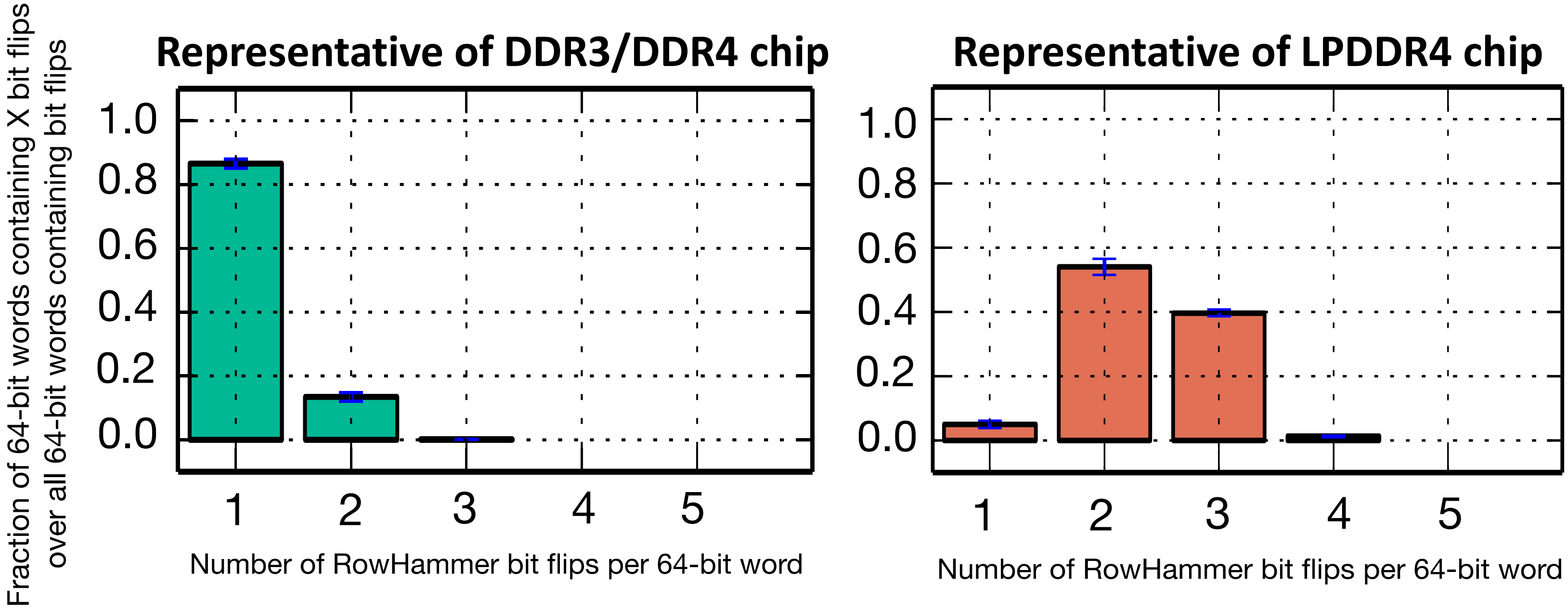
[More analysis in the paper]

SAFARI

4. Spatial Distribution of Bit Flips

Q. How are RowHammer bit flips spatially distributed across a chip?

We normalize data by inducing a bit flip rate of **10^{-6}** in each chip

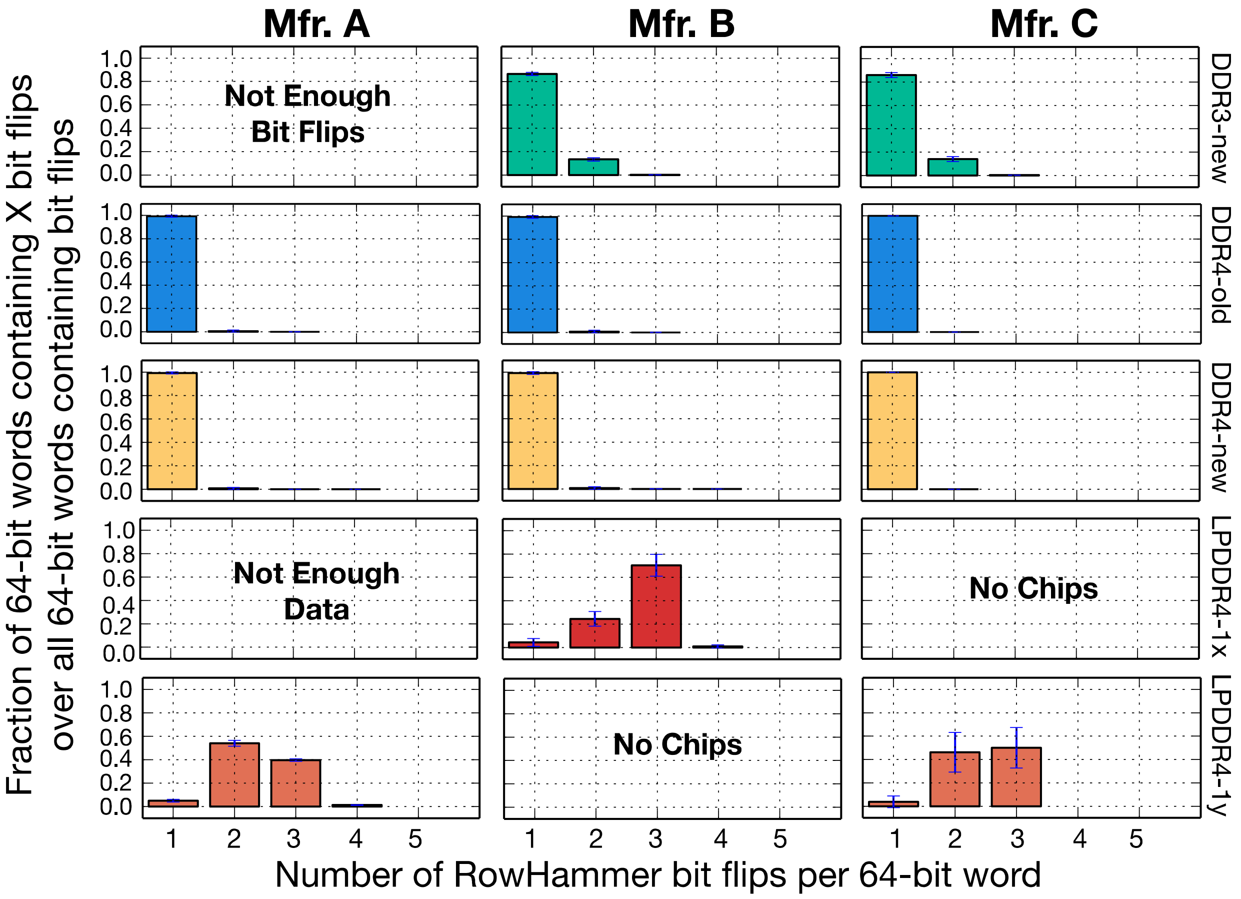


The distribution of RowHammer bit flip density per word **changes significantly in LPDDR4 chips** from other DRAM types

At a bit flip rate of 10^{-6}, a 64-bit word can contain up to **4 bit flips**.
Even at this very low bit flip rate, a **very strong ECC** is required

4. Spatial Distribution of Bit Flips

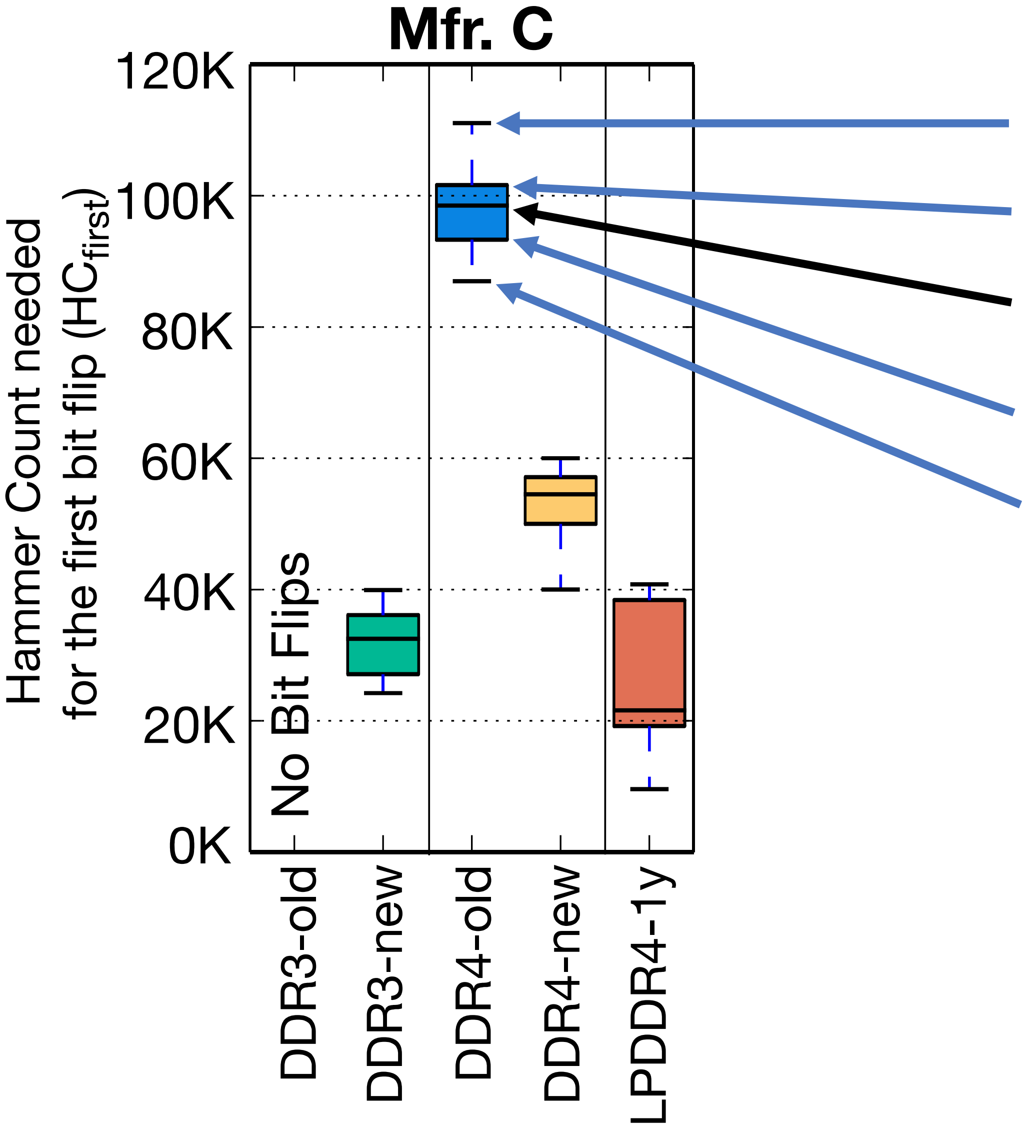
We plot this data for each DRAM type-node configuration per manufacturer



[More analysis in the paper]

5. First RowHammer Bit Flips per Chip

What is the minimum Hammer Count required to cause bit flips (HC_{first})?



Whisker

Q3: 75% point

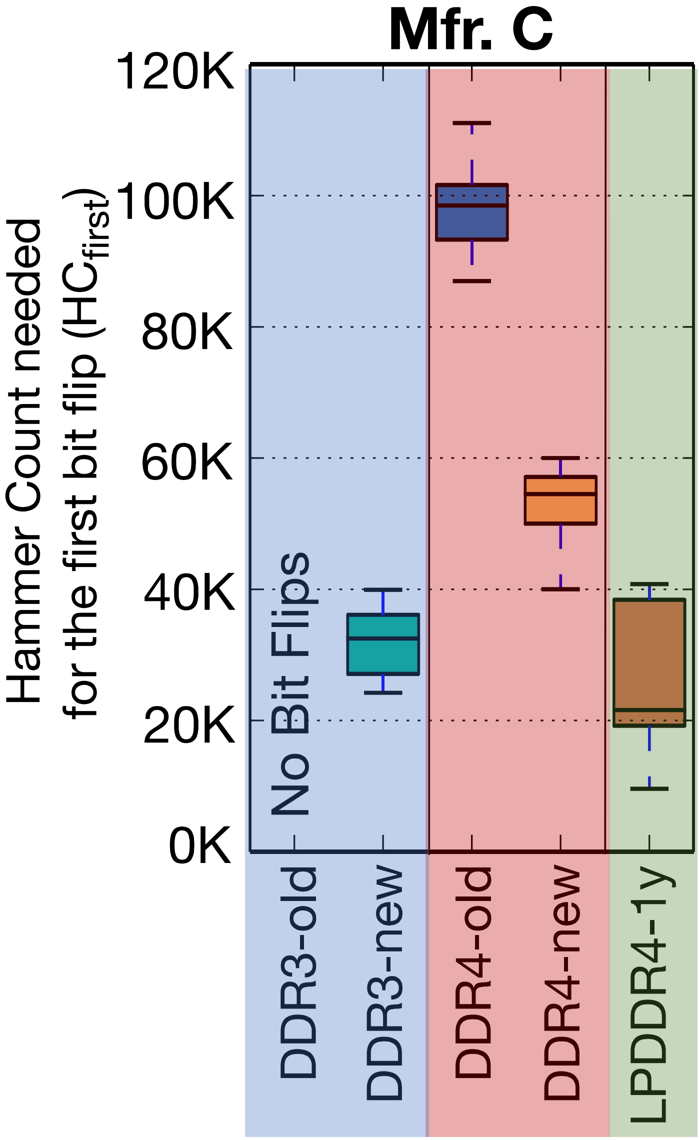
Median: 50%

Q1: 25% point

Whisker

5. First RowHammer Bit Flips per Chip

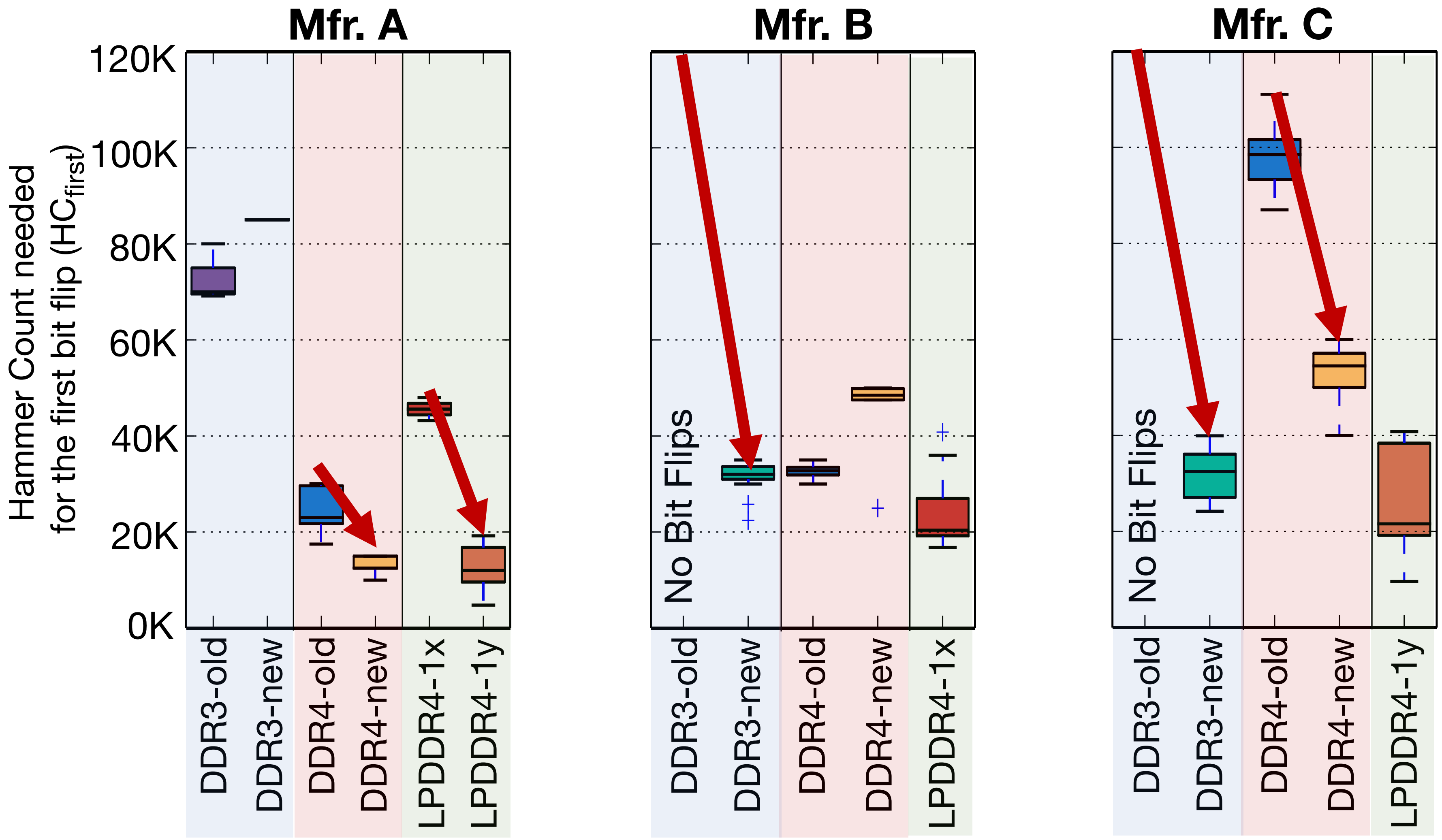
What is the minimum Hammer Count required to cause bit flips (HC_{first})?



We note the different DRAM types on the x-axis: **DDR3**, **DDR4**, **LPDDR4.**

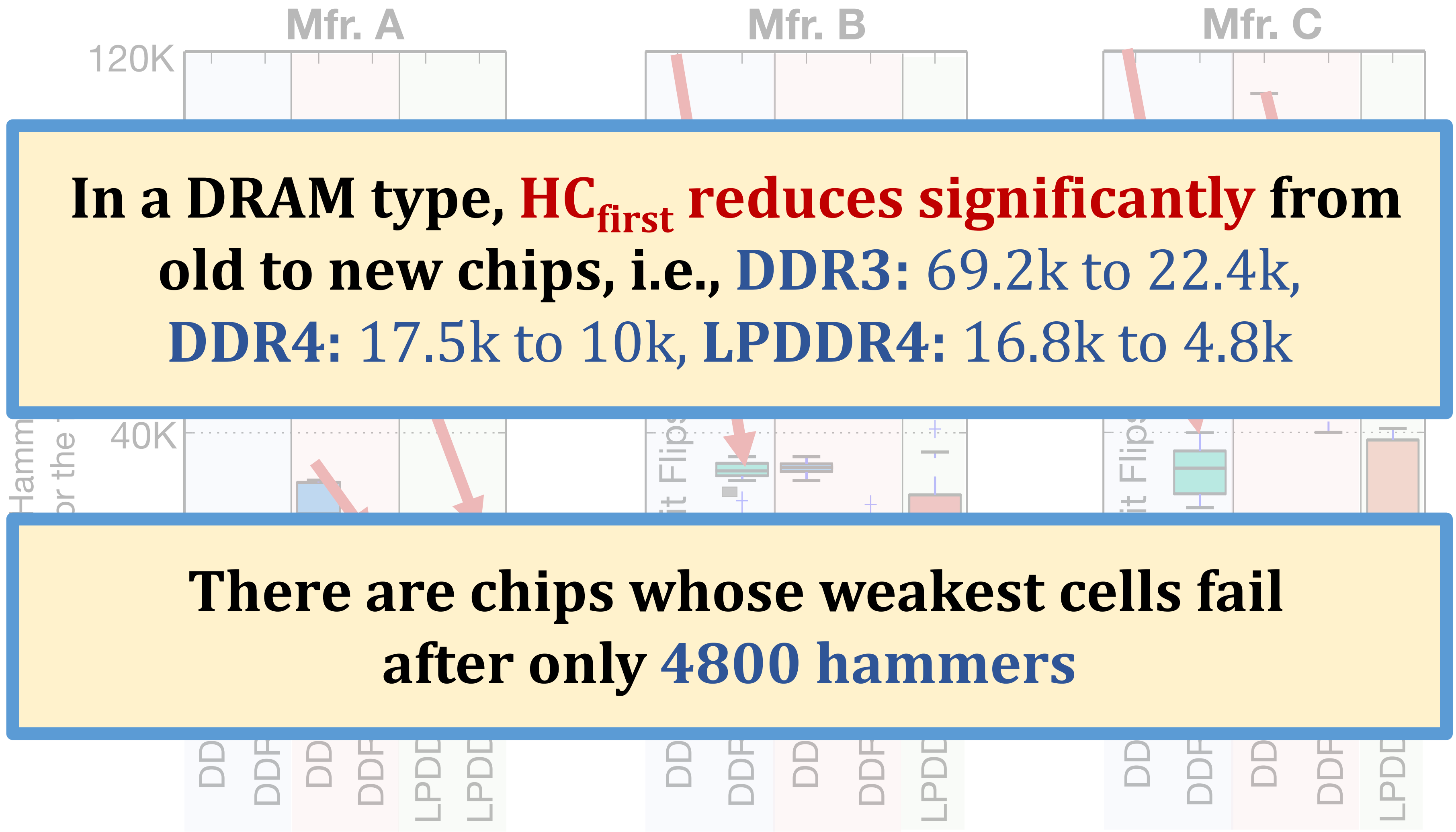
We focus on trends across chips of the same DRAM type to draw conclusions

5. First RowHammer Bit Flips per Chip



Newer chips from a given DRAM manufacturer
more vulnerable to RowHammer

5. First RowHammer Bit Flips per Chip



In a DRAM type, HC_{first} reduces significantly from old to new chips, i.e., DDR3: 69.2k to 22.4k, DDR4: 17.5k to 10k, LPDDR4: 16.8k to 4.8k

There are chips whose weakest cells fail after only 4800 hammers

Newer chips from a given DRAM manufacturer **more** vulnerable to RowHammer

Key Takeaways from 1580 Chips

- Chips of newer DRAM technology nodes are **more vulnerable** to RowHammer

- There are chips today whose weakest cells fail after **only 4800 hammers**

- Chips of newer DRAM technology nodes can exhibit RowHammer bit flips 1) in **more rows** and 2) **farther away** from the victim row.

Evaluation Methodology

- **Cycle-level simulator:** Ramulator [Kim+, CAL'15]
  https://github.com/CMU-SAFARI/ramulator
  - 4GHz, 4-wide, 128 entry instruction window
  - 48 8-core workload mixes randomly drawn from SPEC CPU2006 **(10 < MPKI < 740)**

- **Metrics to evaluate mitigation mechanisms**
  1. ***DRAM Bandwidth Overhead:*** fraction of total system DRAM bandwidth consumption from mitigation mechanism
  2. ***Normalized System Performance:*** normalized weighted speedup to a 100% baseline

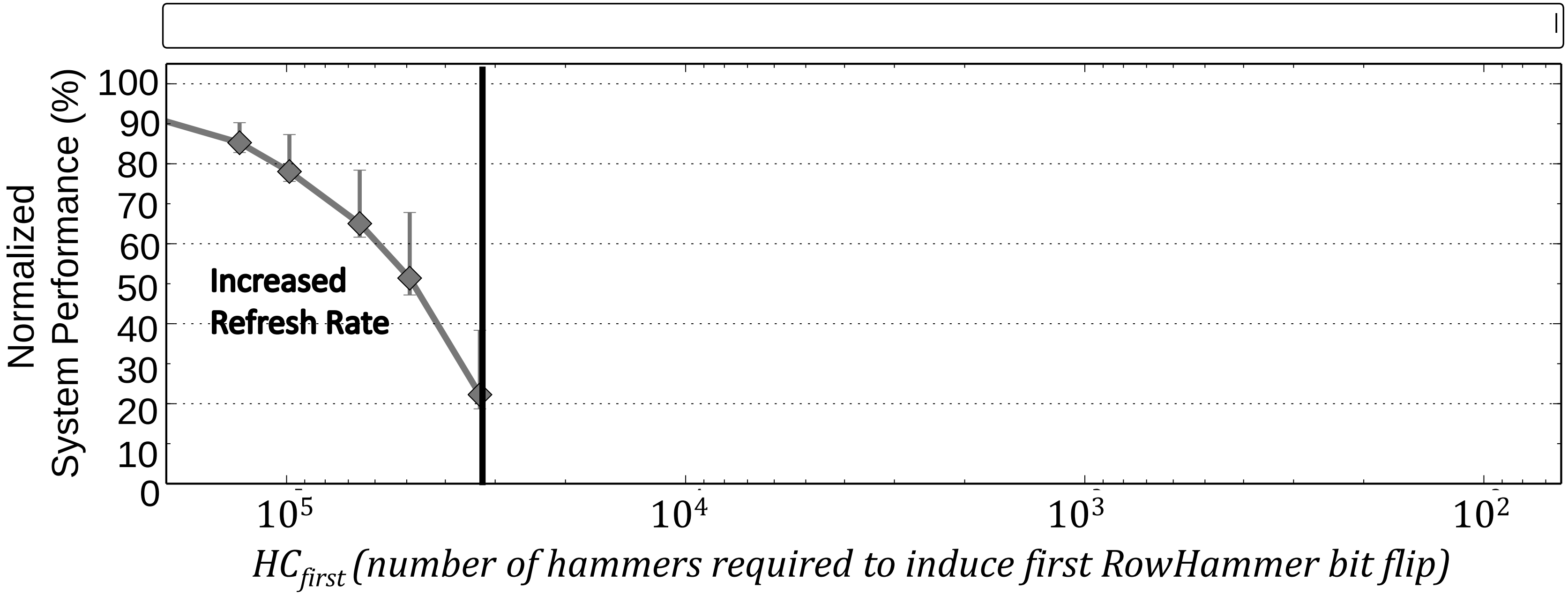
Evaluation Methodology

- We evaluate **five** state-of-the-art mitigation mechanisms:
  - **Increased Refresh Rate** **[Kim+, ISCA'14]**
  - **PARA** **[Kim+, ISCA'14]**
  - **ProHIT** **[Son+, DAC'17]**
  - **MRLoc** **[You+, DAC'19]**
  - **TWiCe** **[Lee+, ISCA'19]**

- and **one** ideal refresh-based mitigation mechanism:
  - **Ideal**

- **More detailed descriptions in the paper on:**
  - Descriptions of mechanisms in our paper and the original publications
  - How we scale each mechanism to more vulnerable DRAM chips (lower $\mathbf{HC_{first}}$)

Mitigation Mech. Eval. (Increased Refresh)

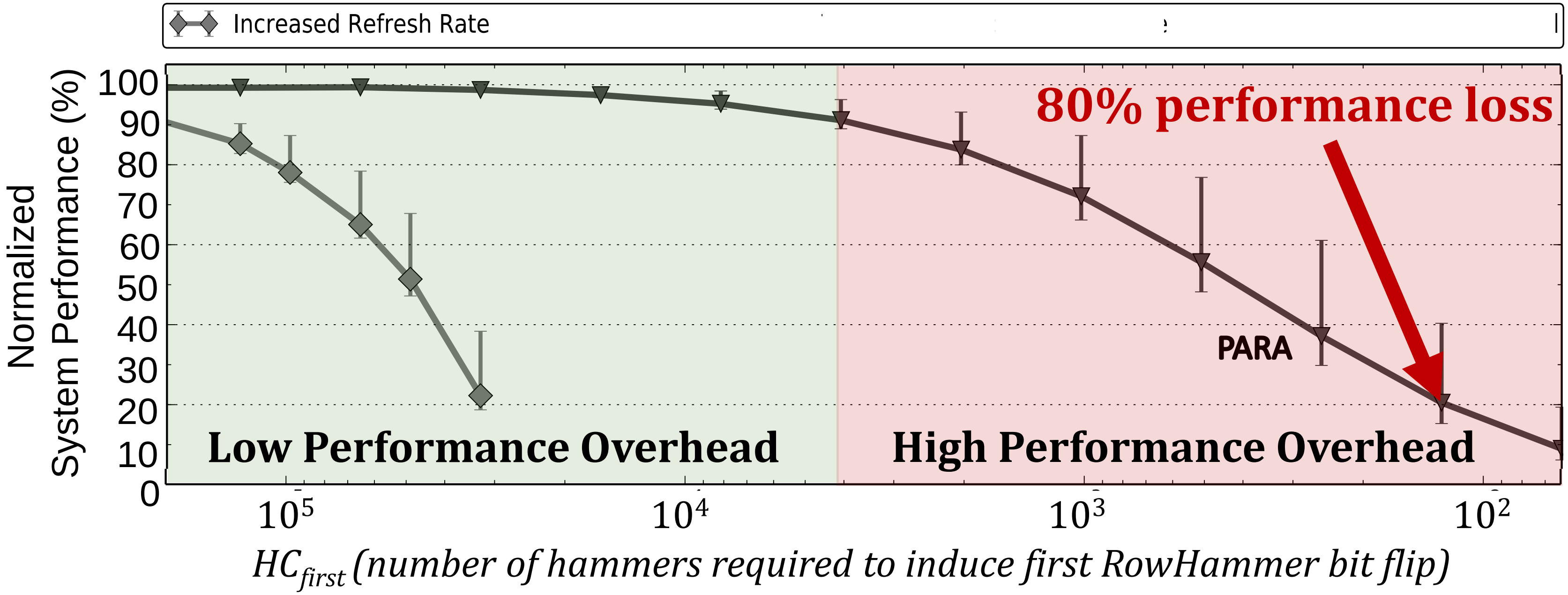


Substantial overhead for high HC_{first} values.

This mechanism does not support HC_{first} < 32k due to the prohibitively high refresh rates required

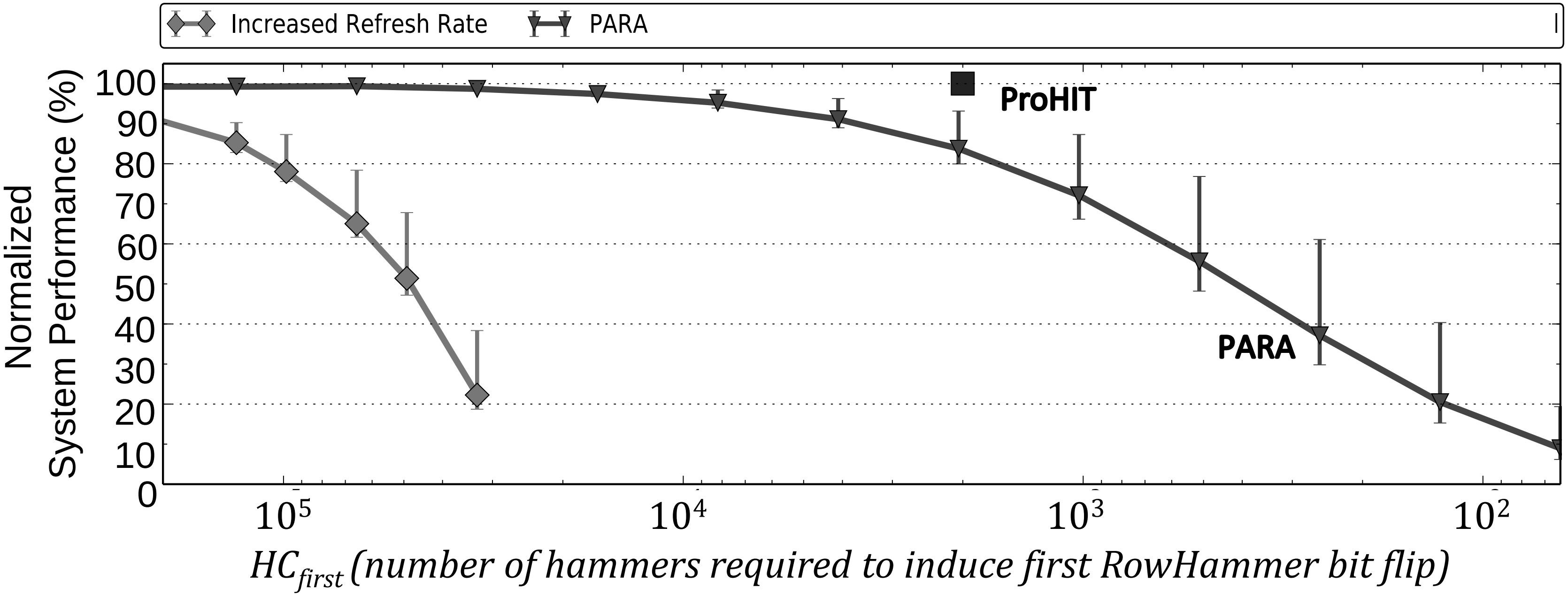
SAFARI

Mitigation Mechanism Evaluation (PARA)

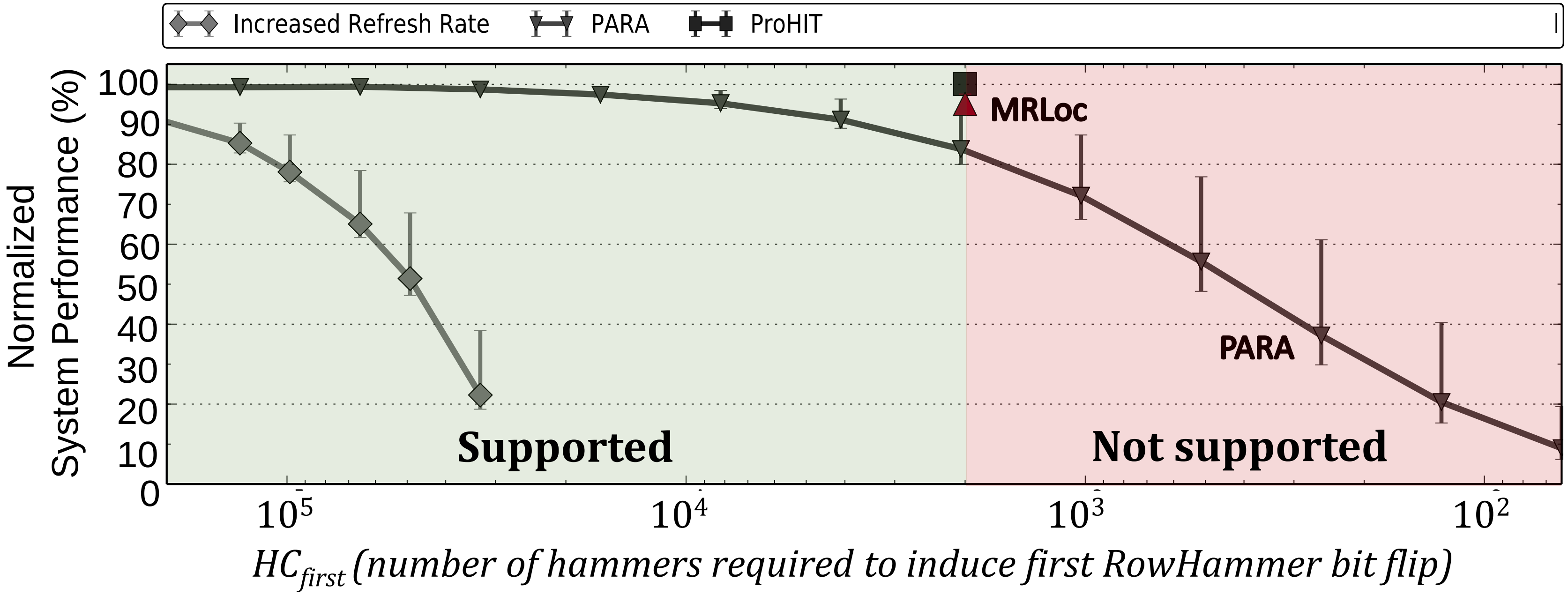


80% performance loss

Mitigation Mechanism Evaluation (ProHIT)

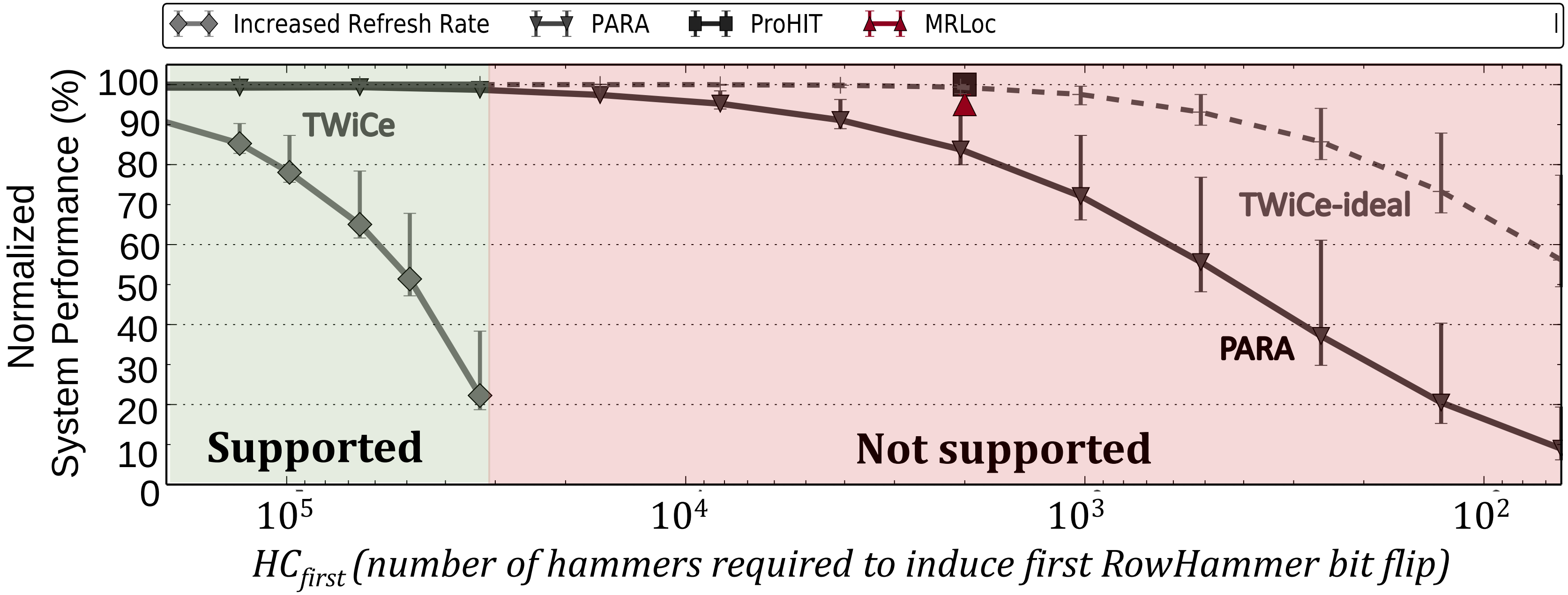


Mitigation Mechanism Evaluation (MRLoc)



Models for scaling ProHIT and MRLoc for HC_{first} < 2k are not provided and how to do so is not intuitive

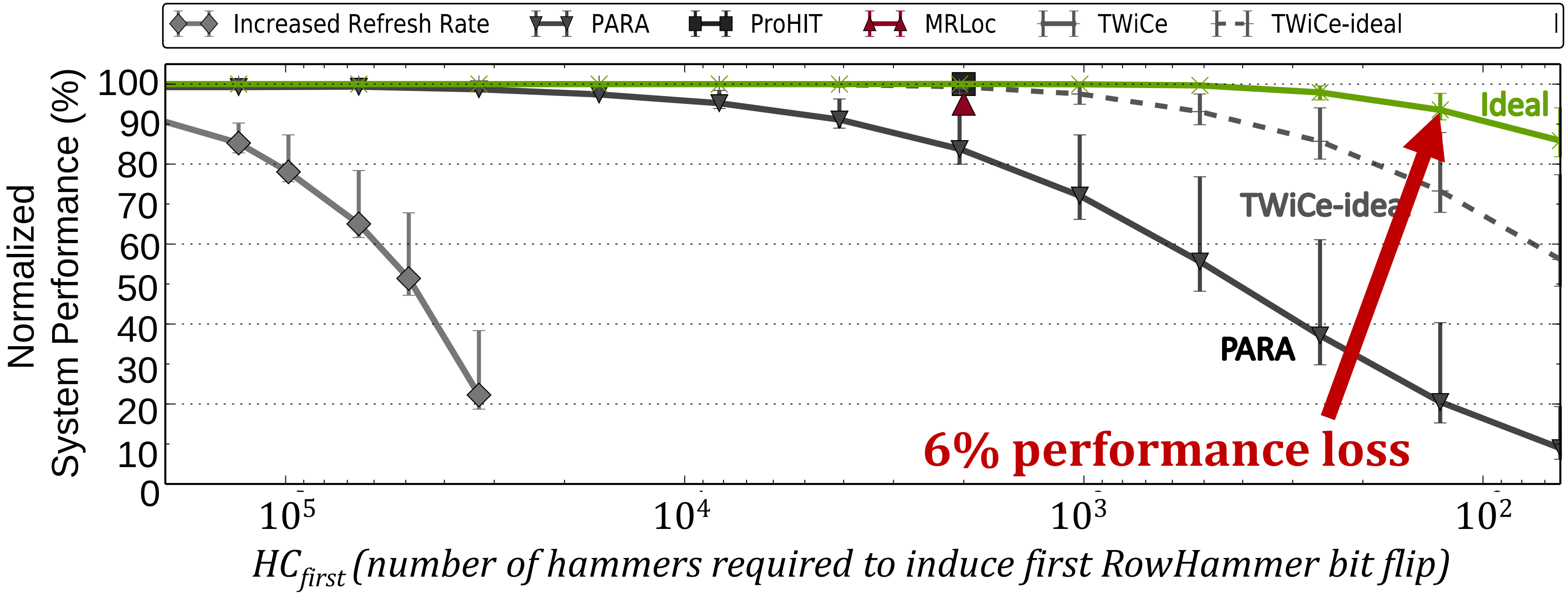
Mitigation Mechanism Evaluation (TWiCe)



TWiCe does not support HC_{first} < 32k.

We evaluate an ideal scalable version (TWiCe-ideal) assuming it solves two critical design issues

Mitigation Mechanism Evaluation (Ideal)

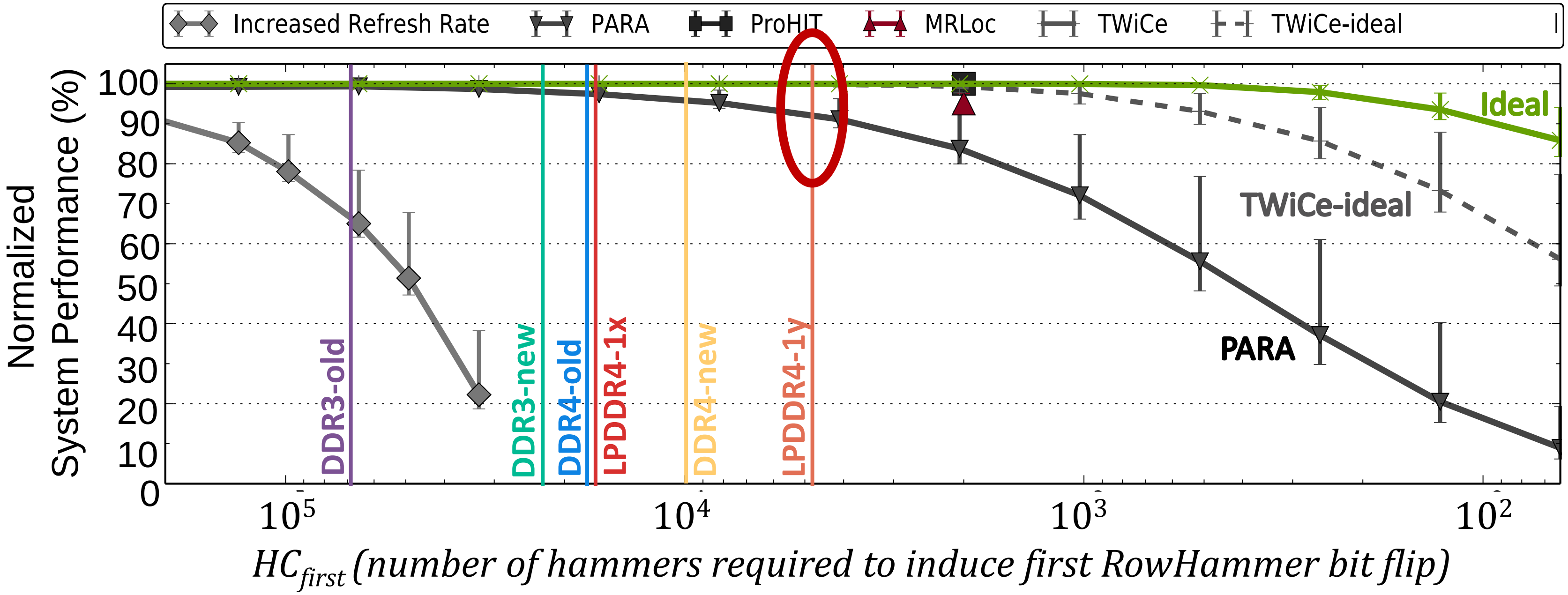


HC_{first} (number of hammers required to induce first RowHammer bit flip)

6% performance loss

Ideal mechanism issues a refresh command to a row only right before the row can potentially experience a RowHammer bit flip

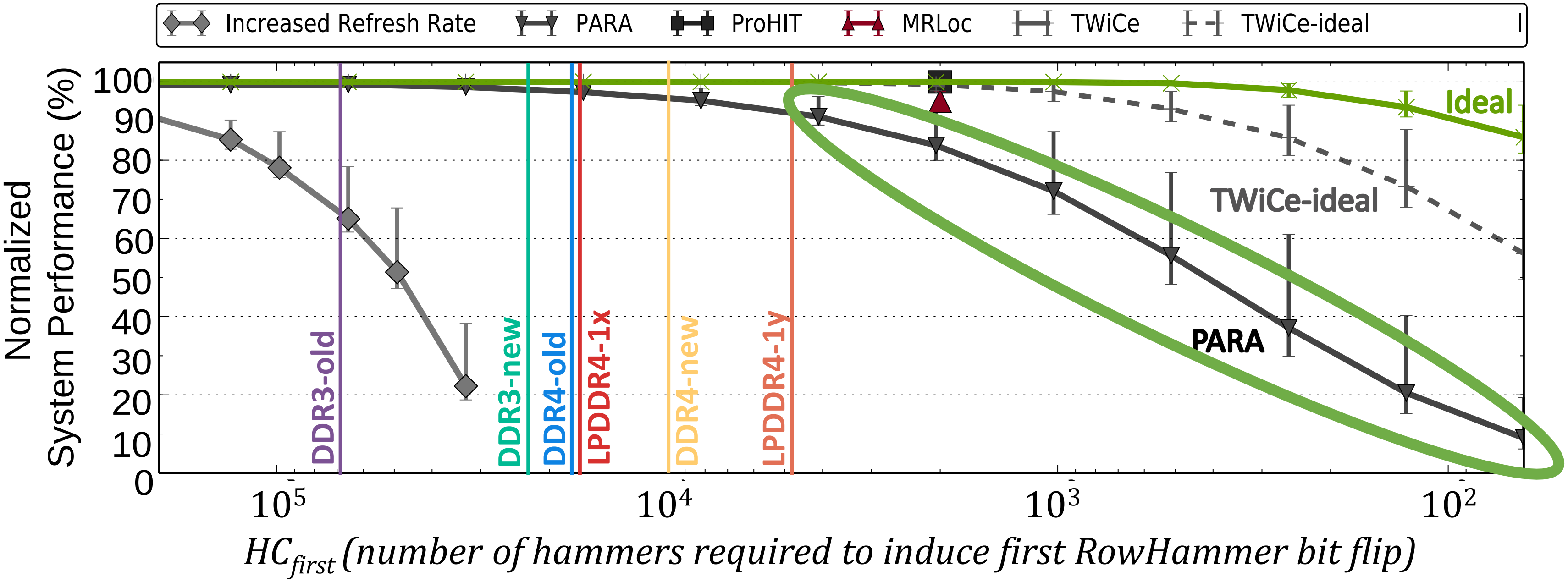
Mitigation Mechanism Evaluation



PARA, ProHIT, and MRLoc mitigate RowHammer bit flips in worst chips today with reasonable system performance (92%, 100%, 100%)

SAFARI

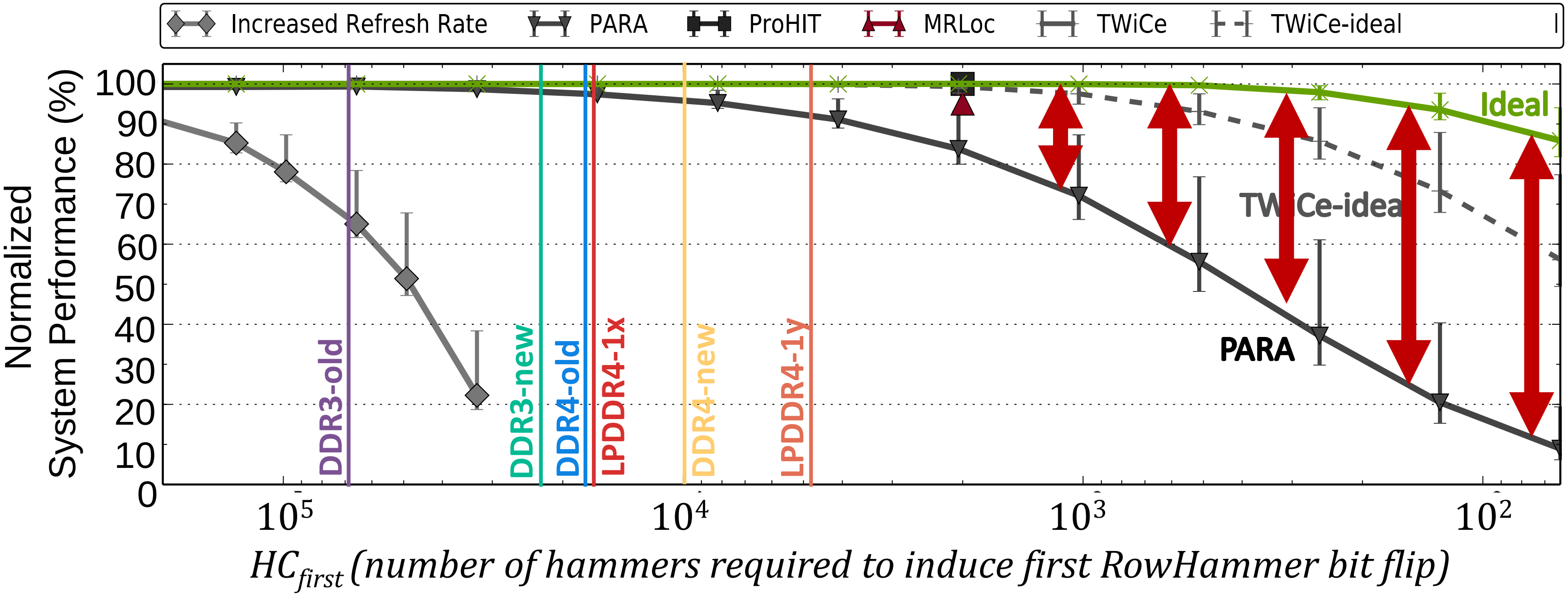
Mitigation Mechanism Evaluation



Only PARA's design scales to low HC_{first} values
but has very low normalized system performance

SAFARI

Mitigation Mechanism Evaluation



Ideal mechanism is significantly better than any existing mechanism for HC_{first} < 1024

Significant opportunity for developing a RowHammer solution with low performance overhead that supports low HC_{first}

Key Takeaways from Mitigation Mechanisms

- Existing RowHammer mitigation mechanisms can prevent RowHammer attacks with **reasonable system performance overhead** in DRAM chips today

- Existing RowHammer mitigation mechanisms **do not scale well** to DRAM chips more vulnerable to RowHammer

- There is still **significant opportunity** for developing a mechanism that is **scalable with low overhead**

Additional Details in the Paper

- **Single-cell RowHammer bit flip probability**
- More details on our **data pattern dependence** study
- Analysis of **Error Correcting Codes (ECC)** in mitigating RowHammer bit flips
- Additional **observations** on our data
- **Methodology details** for characterizing DRAM
- Further discussion on comparing data across different infrastructures
- **Discussion on scaling** each mitigation mechanism

RowHammer Solutions Going Forward

Two promising directions for new RowHammer solutions:

1. **DRAM-system cooperation**
   - We believe the DRAM and system should cooperate more to provide a **holistic** solution can prevent RowHammer at **low cost**

2. **Profile-guided**
   - Accurate **profile of RowHammer-susceptible cells** in DRAM provides a powerful substrate for building **targeted** RowHammer solutions, e.g.:
     - Only increase the refresh rate for rows containing RowHammer-susceptible cells
   - A **fast and accurate** profiling mechanism is a key research challenge for developing low-overhead and scalable RowHammer solutions

Conclusion

- We characterized **1580 DRAM** chips of different DRAM types, technology nodes, and manufacturers.

- We studied **five** state-of-the-art RowHammer mitigation mechanisms and an ideal refresh-based mechanism

- We made **two key observations**
  1. **RowHammer is getting much worse**. It takes much fewer hammers to induce RowHammer bit flips in newer chips
     - e.g., **DDR3:** 69.2k to 22.4k, **DDR4:** 17.5k to 10k, **LPDDR4:** 16.8k to 4.8k
  2. **Existing mitigation mechanisms do not scale** to DRAM chips that are more vulnerable to RowHammer
     - e.g., 80% performance loss when the hammer count to induce the first bit flip is 128

- We **conclude** that it is **critical** to do more research on RowHammer and develop scalable mitigation mechanisms to prevent RowHammer in future systems

Revisiting RowHammer

An Experimental Analysis of Modern Devices and Mitigation Techniques

Jeremie S. Kim **Minesh Patel**

A. Giray Yağlıkçı **Hasan Hassan**

Roknoddin Azizi **Lois Orosa** **Onur Mutlu**

SAFARI

ETH Zürich

Carnegie Mellon

Revisiting RowHammer in 2020 (I)

- Jeremie S. Kim, Minesh Patel, A. Giray Yaglikci, Hasan Hassan, Roknoddin Azizi, Lois Orosa, and Onur Mutlu,
  "Revisiting RowHammer: An Experimental Analysis of Modern Devices and Mitigation Techniques"
  *Proceedings of the 47th International Symposium on Computer Architecture (**ISCA**)*, Valencia, Spain, June 2020.
  [Slides (pptx) (pdf)]
  [Lightning Talk Slides (pptx) (pdf)]
  [Talk Video (20 minutes)]
  [Lightning Talk Video (3 minutes)]

Revisiting RowHammer: An Experimental Analysis of Modern DRAM Devices and Mitigation Techniques

Jeremie S. Kim[§†] Minesh Patel[§] A. Giray Yağlıkçı[§]
Hasan Hassan[§] Roknoddin Azizi[§] Lois Orosa[§] Onur Mutlu[§†]

[§]*ETH Zürich* [†]*Carnegie Mellon University*

Future Memory Reliability/Security Challenges

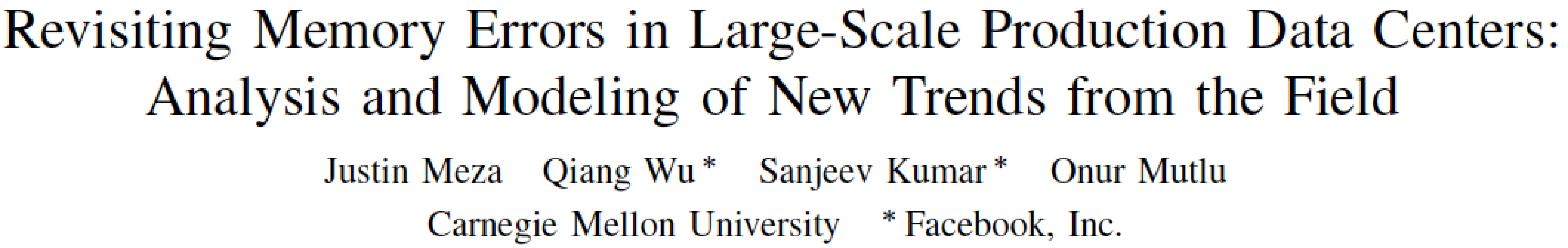
Future of Main Memory

- DRAM is becoming less reliable → more vulnerable

Large-Scale Failure Analysis of DRAM Chips

- Analysis and modeling of memory errors found in all of Facebook’s server fleet

- Justin Meza, Qiang Wu, Sanjeev Kumar, and Onur Mutlu, **"Revisiting Memory Errors in Large-Scale Production Data Centers: Analysis and Modeling of New Trends from the Field"** *Proceedings of the 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (**DSN**)*, Rio de Janeiro, Brazil, June 2015.
  [Slides (pptx) (pdf)] [DRAM Error Model]

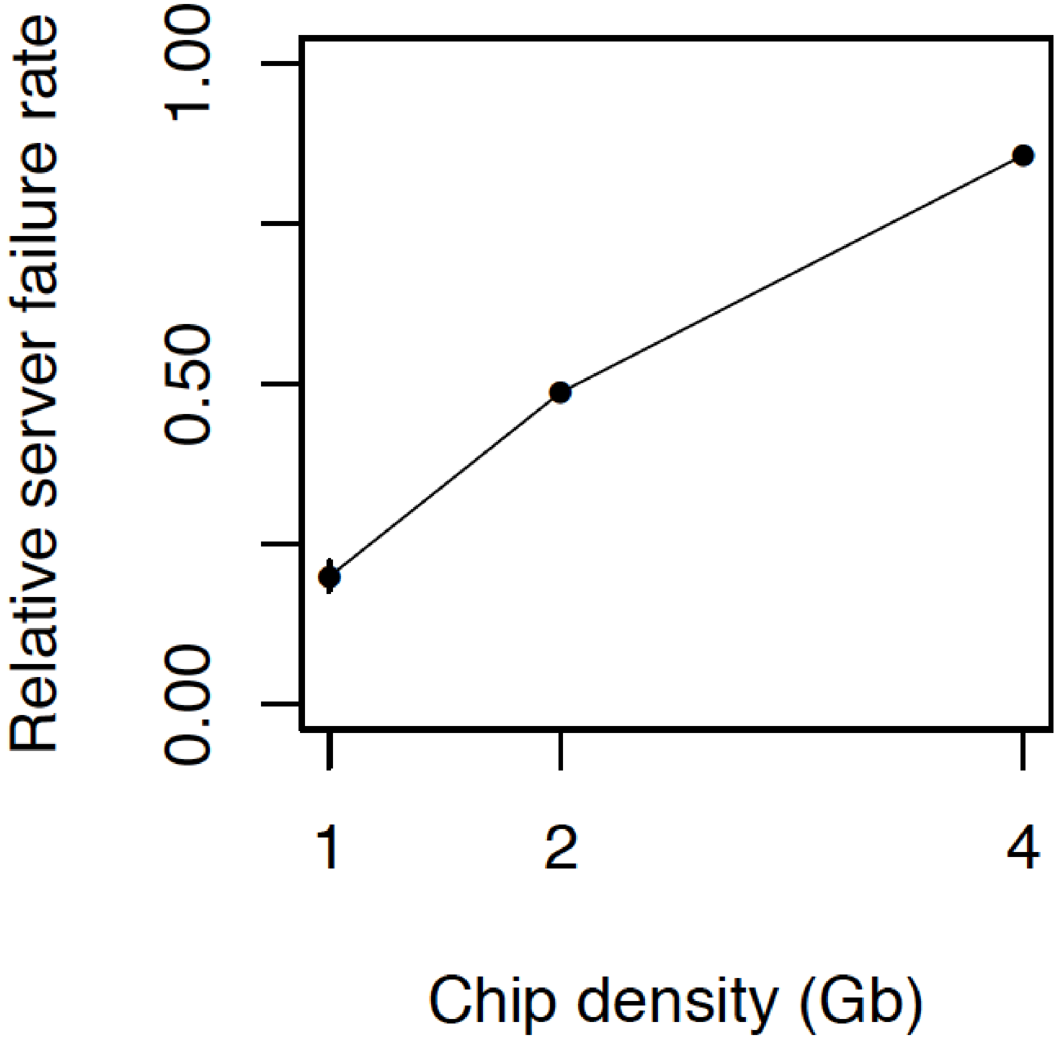


Revisiting Memory Errors in Large-Scale Production Data Centers: Analysis and Modeling of New Trends from the Field

Justin Meza Qiang Wu * Sanjeev Kumar * Onur Mutlu

Carnegie Mellon University * Facebook, Inc.

DRAM Reliability Reducing



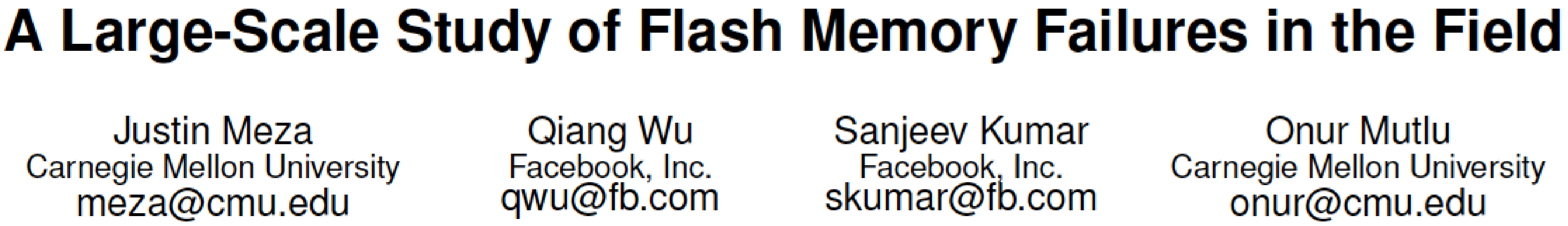
Intuition: quadratic increase in capacity

Meza+, “Revisiting Memory Errors in Large-Scale Production Data Centers,” DSN’15.

Aside: SSD Error Analysis in the Field

- 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]



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

Future of Main Memory

- DRAM is becoming less reliable → more vulnerable

- Due to difficulties in DRAM scaling, other problems may also appear (or they may be going unnoticed)

- Some errors may already be slipping into the field
  - Read disturb errors (Rowhammer)
  - Retention errors
  - Read errors, write errors
  - …

- These errors can also pose security vulnerabilities

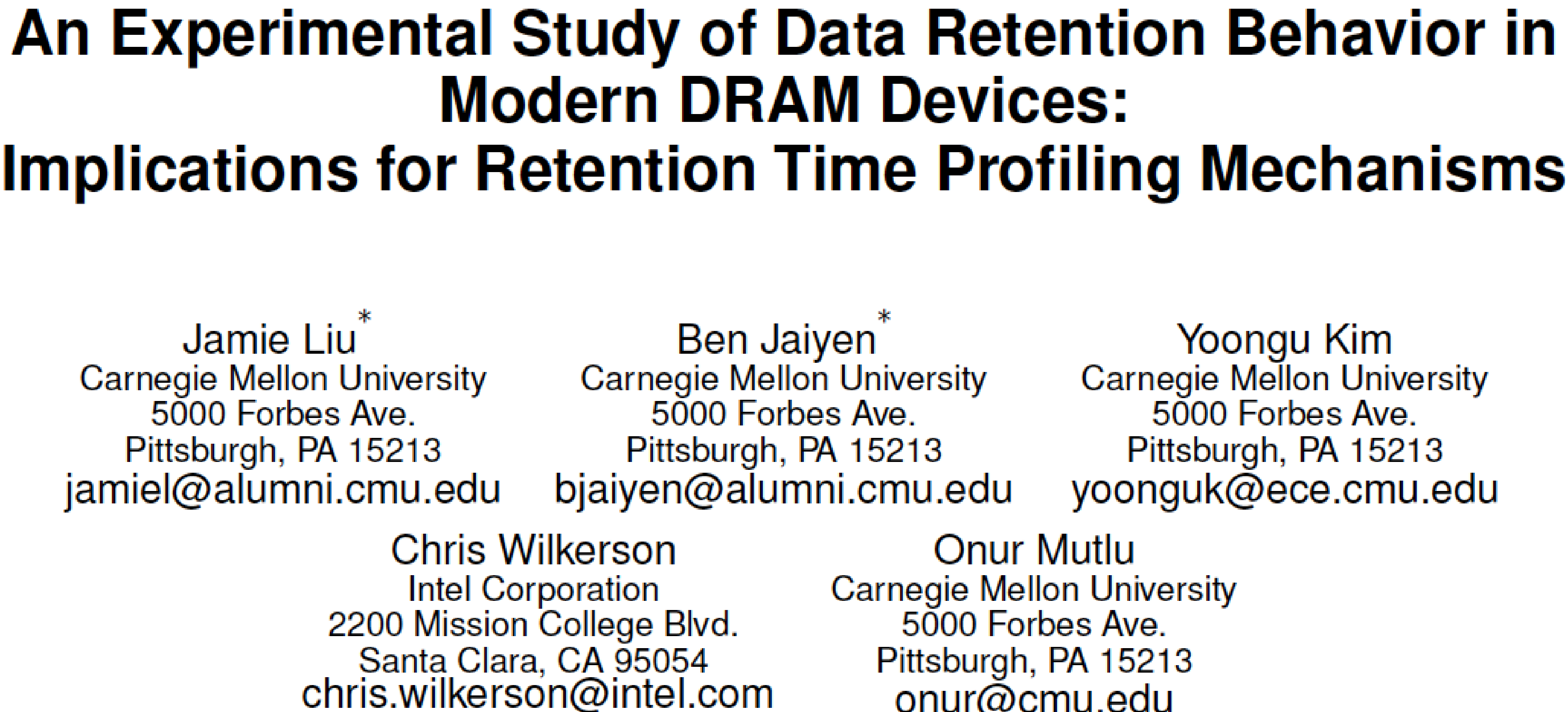
DRAM Data Retention Time Failures

- Determining the data retention time of a cell/row is getting more difficult

- Retention failures may already be slipping into the field

Analysis of Data Retention Failures [ISCA'13]

- Jamie Liu, Ben Jaiyen, Yoongu Kim, Chris Wilkerson, and Onur Mutlu,
  "An Experimental Study of Data Retention Behavior in Modern DRAM Devices: Implications for Retention Time Profiling Mechanisms"
  *Proceedings of the 40th International Symposium on Computer Architecture (**ISCA**)*, Tel-Aviv, Israel, June 2013. Slides (ppt) Slides (pdf)



An Experimental Study of Data Retention Behavior in Modern DRAM Devices: Implications for Retention Time Profiling Mechanisms

Jamie Liu[*]
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
jamiel@alumni.cmu.edu

Ben Jaiyen[*]
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
bjaiyen@alumni.cmu.edu

Yoongu Kim
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
yoonguk@ece.cmu.edu

Chris Wilkerson
Intel Corporation
2200 Mission College Blvd.
Santa Clara, CA 95054
chris.wilkerson@intel.com

Onur Mutlu
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
onur@cmu.edu

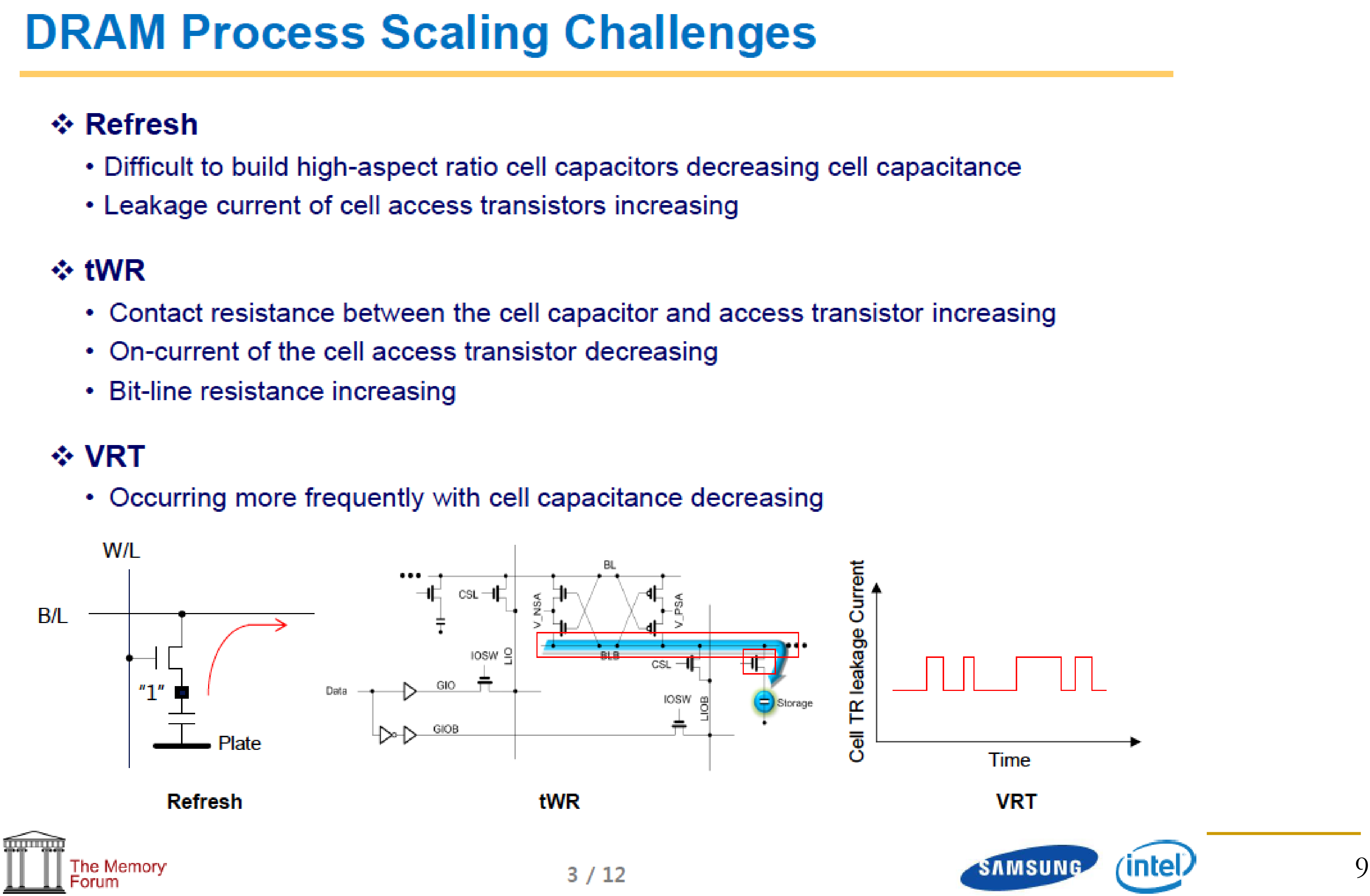
Two Challenges to Retention Time Profiling

- Data Pattern Dependence (DPD) of retention time

- Variable Retention Time (VRT) phenomenon

https://www.youtube.com/watch?v=v702wUnaWGE

Industry Is Writing Papers About It, Too



DRAM Process Scaling Challenges

❖ **Refresh**

- Difficult to build high-aspect ratio cell capacitors decreasing cell capacitance
- Leakage current of cell access transistors increasing

❖ **tWR**

- Contact resistance between the cell capacitor and access transistor increasing
- On-current of the cell access transistor decreasing
- Bit-line resistance increasing

❖ **VRT**

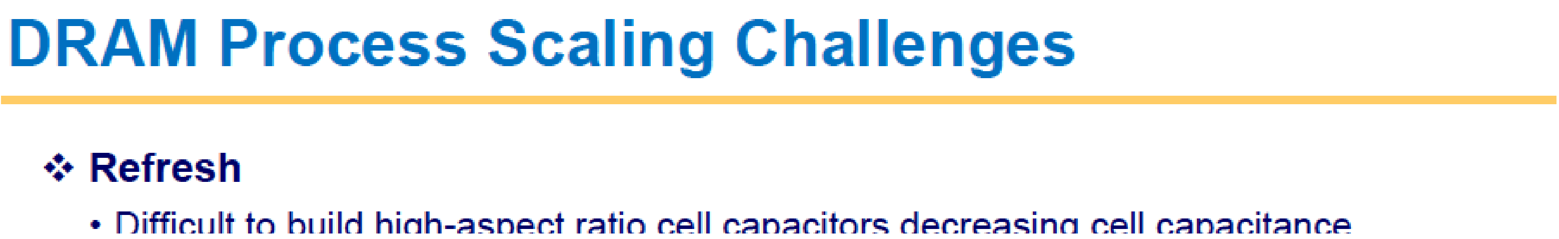
- Occurring more frequently with cell capacitance decreasing

Refresh **tWR** **VRT**

The Memory Forum

3 / 12

Industry Is Writing Papers About It, Too

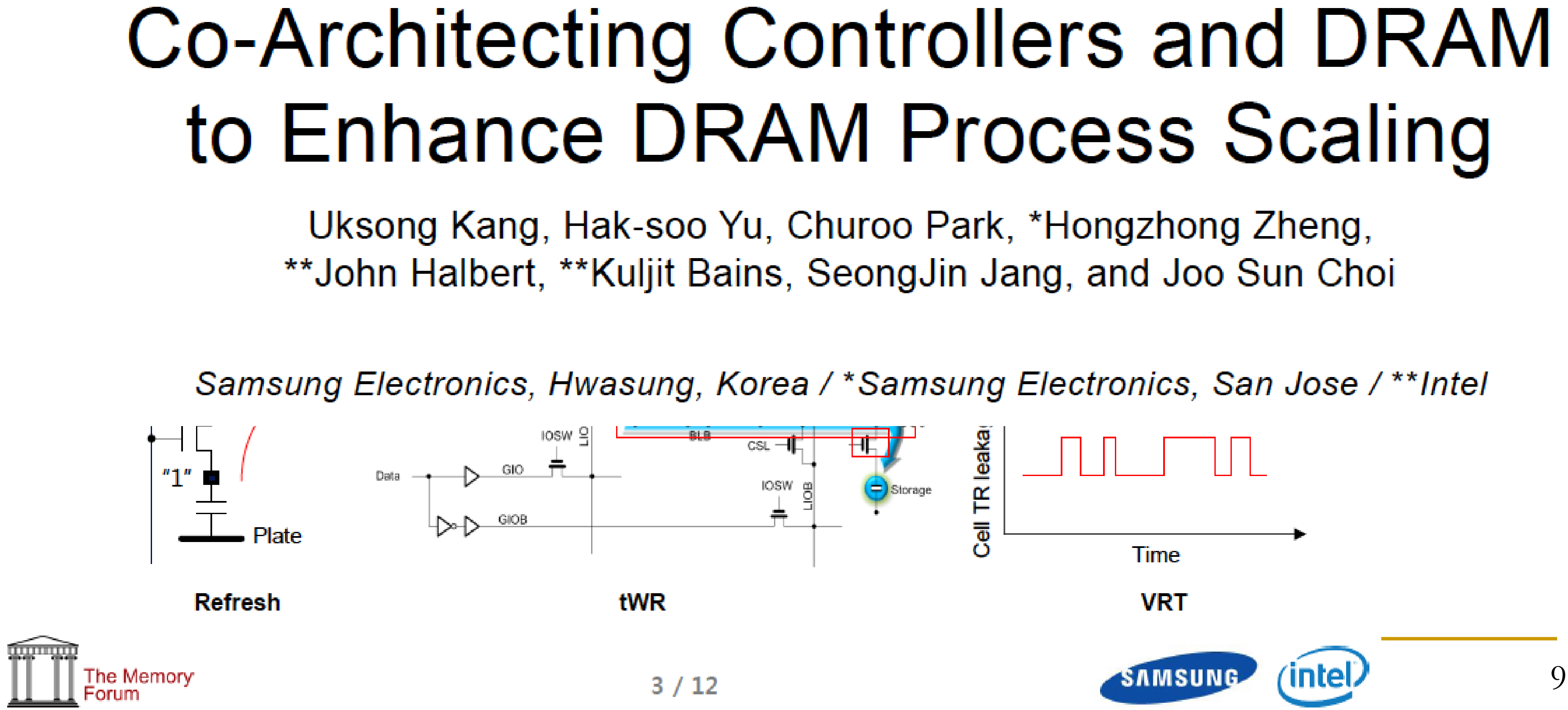


DRAM Process Scaling Challenges

❖ **Refresh**

- Difficult to build high-aspect ratio cell capacitors decreasing cell capacitance

THE MEMORY FORUM 2014



Co-Architecting Controllers and DRAM to Enhance DRAM Process Scaling

Uksong Kang, Hak-soo Yu, Churoo Park, *Hongzhong Zheng, **John Halbert, **Kuljit Bains, SeongJin Jang, and Joo Sun Choi

*Samsung Electronics, Hwasung, Korea / *Samsung Electronics, San Jose / **Intel*

Refresh **tWR** **VRT**

3 / 12

Keeping Future Memory Secure

How Do We Keep Memory Secure?

- DRAM
- Flash memory
- Emerging Technologies
  - Phase Change Memory
  - STT-MRAM
  - RRAM, memristors
  - …

Many Errors and Their Mitigation [PIEEE'17]

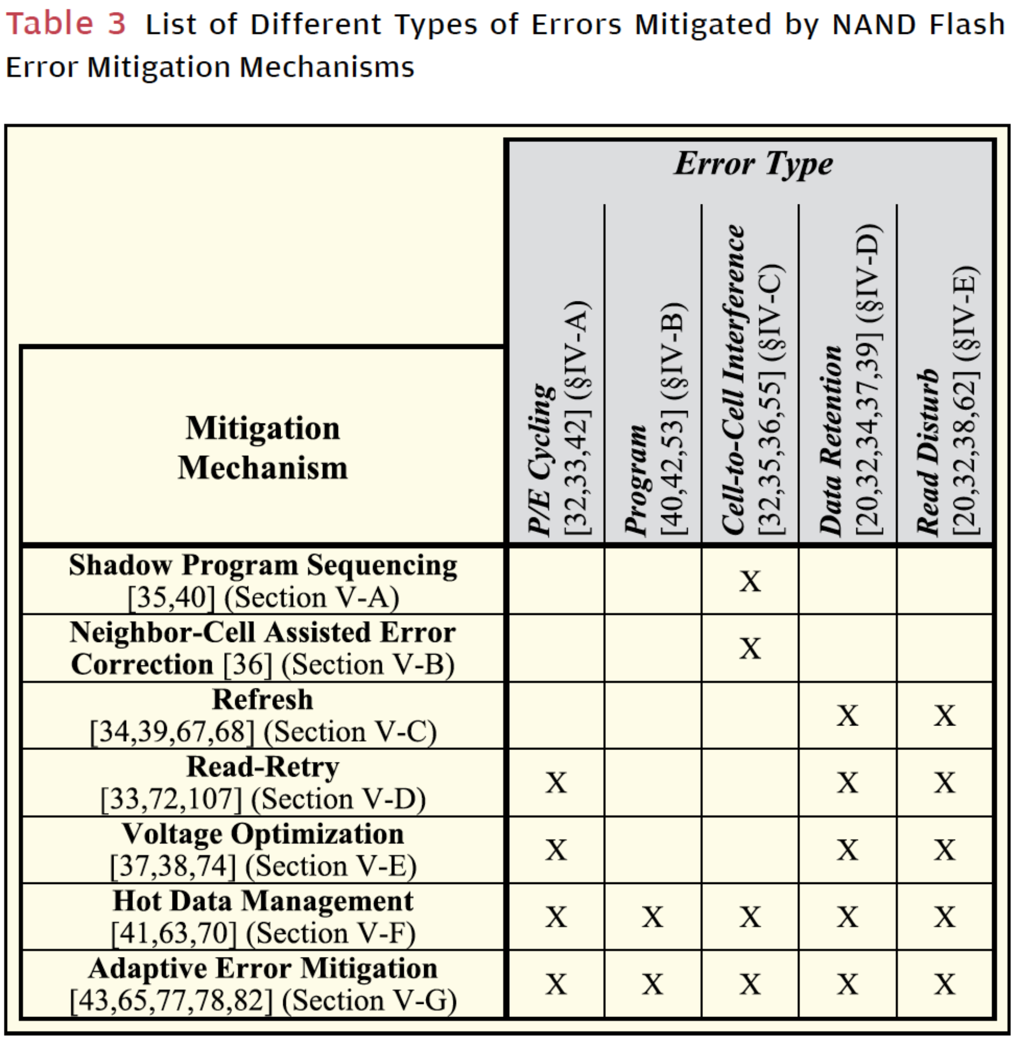


Table 3 List of Different Types of Errors Mitigated by NAND Flash Error Mitigation Mechanisms

	Error Type				
Mitigation Mechanism	***P/E Cycling*** [32,33,42] (§IV-A)	***Program*** [40,42,53] (§IV-B)	***Cell-to-Cell Interference*** [32,35,36,55] (§IV-C)	***Data Retention*** [20,32,34,37,39] (§IV-D)	***Read Disturb*** [20,32,38,62] (§IV-E)
Shadow Program Sequencing [35,40] (Section V-A)			X		
Neighbor-Cell Assisted Error Correction [36] (Section V-B)			X		
Refresh [34,39,67,68] (Section V-C)				X	X
Read-Retry [33,72,107] (Section V-D)	X			X	X
Voltage Optimization [37,38,74] (Section V-E)	X			X	X
Hot Data Management [41,63,70] (Section V-F)	X	X	X	X	X
Adaptive Error Mitigation [43,65,77,78,82] (Section V-G)	X	X	X	X	X

Cai+, "Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives," Proc. IEEE 2017.

Solution Direction: Principled Designs

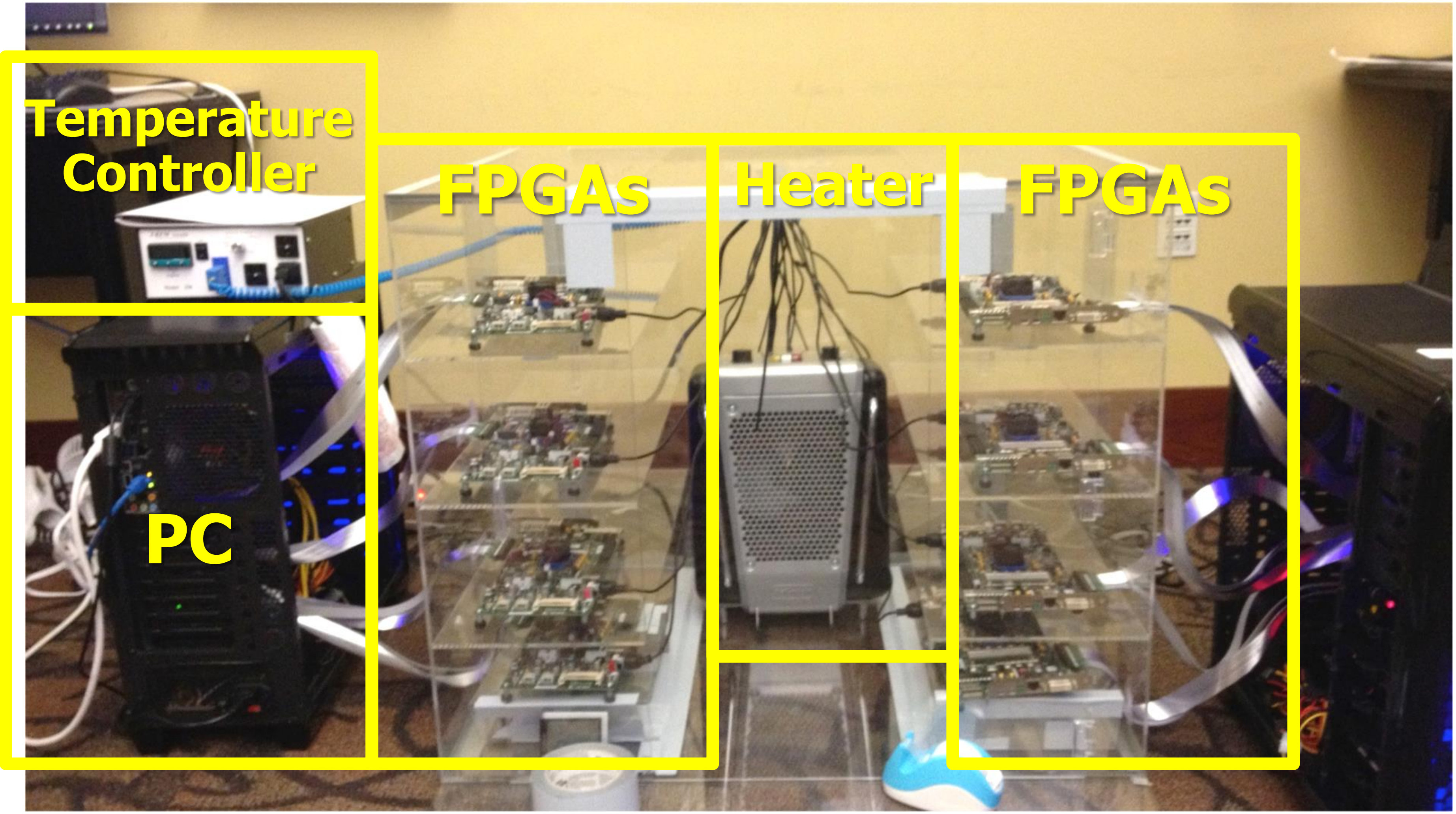
Design fundamentally secure computing architectures

Predict and prevent such safety issues

Architecting Future Memory for Security

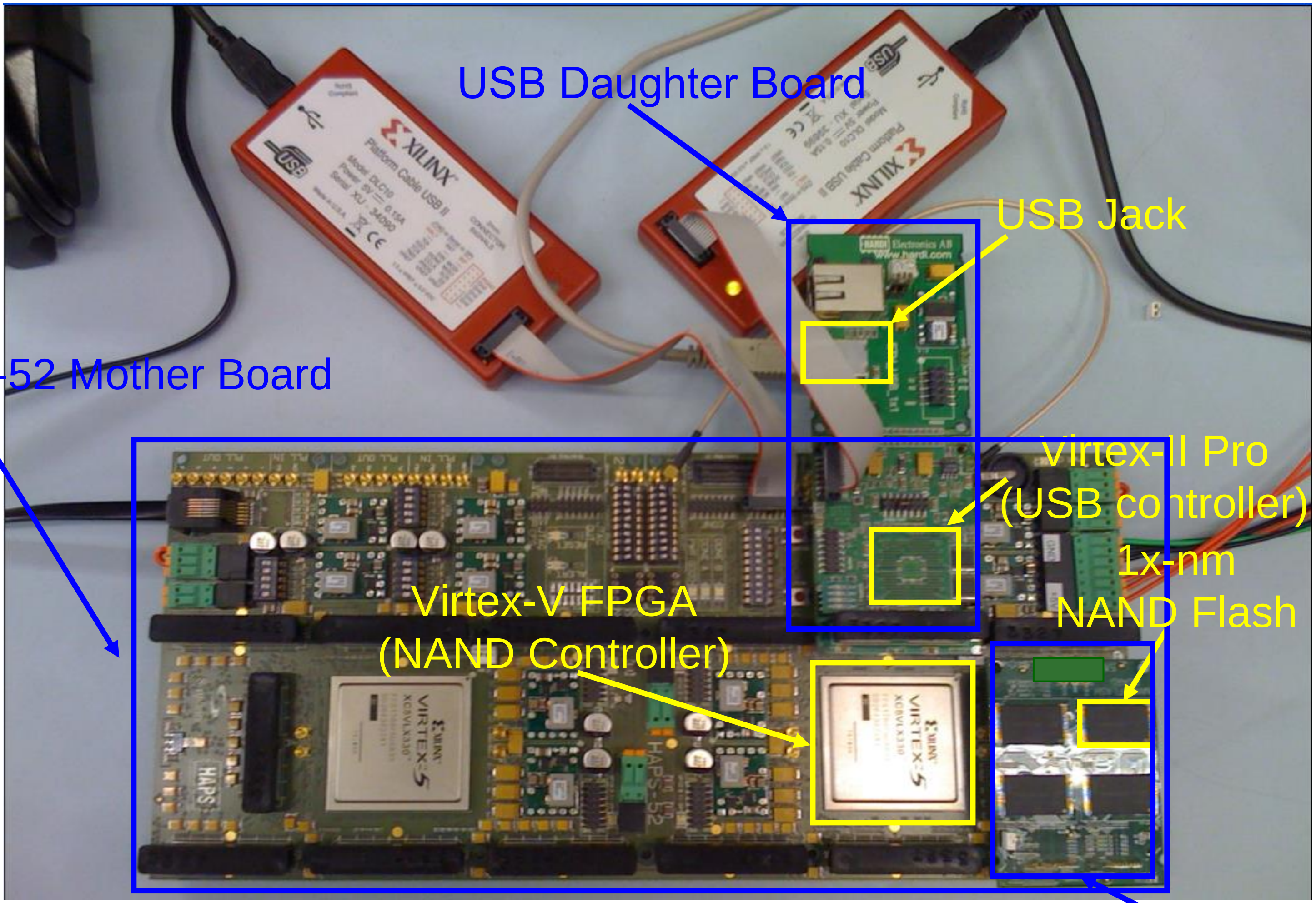
- **Understand**: Methods for vulnerability modeling & discovery
  - Modeling and prediction based on real (device) data and analysis
  - Understanding vulnerabilities
  - Developing reliable metrics
- **Architect**: Principled architectures with security as key concern
  - Good partitioning of duties across the stack
  - Cannot give up performance and efficiency
  - Patch-ability in the field
- **Design & Test**: Principled design, automation, (online) testing
  - Design for security
  - High coverage and good interaction with system reliability methods

Understand and Model with Experiments (DRAM)



Kim+, "Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors," ISCA 2014.

Understand and Model with Experiments (Flash)

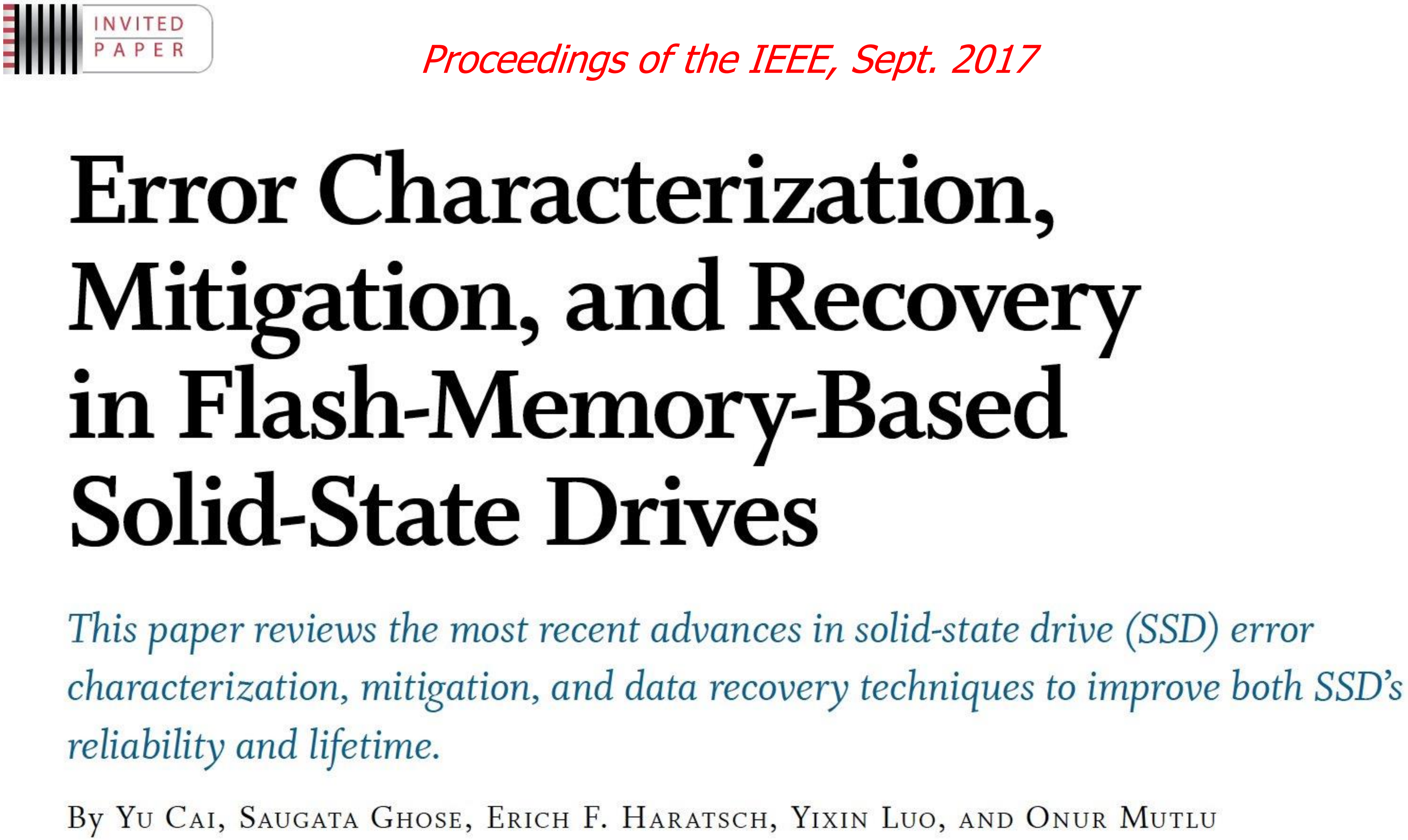


[DATE 2012, ICCD 2012, DATE 2013, ITJ 2013, ICCD 2013, SIGMETRICS 2014, HPCA 2015, DSN 2015, MSST 2015, JSAC 2016, HPCA 2017, DFRWS 2017, PIEEE 2017, HPCA 2018, SIGMETRICS 2018]

Cai+, "Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives," Proc. IEEE 2017.

Understanding Flash Memory Reliability

Proceedings of the IEEE, Sept. 2017



INVITED PAPER

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

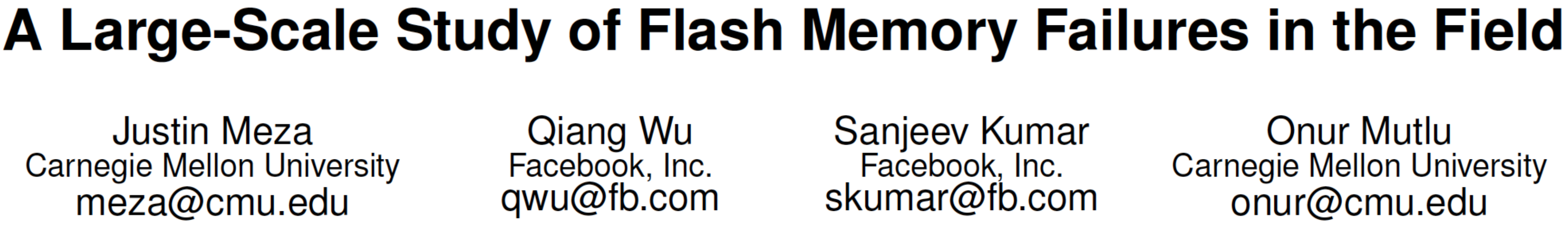
This paper reviews the most recent advances in solid-state drive (SSD) error characterization, mitigation, and data recovery techniques to improve both SSD's reliability and lifetime.

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

https://arxiv.org/pdf/1706.08642

Understanding Flash Memory Reliability

- 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

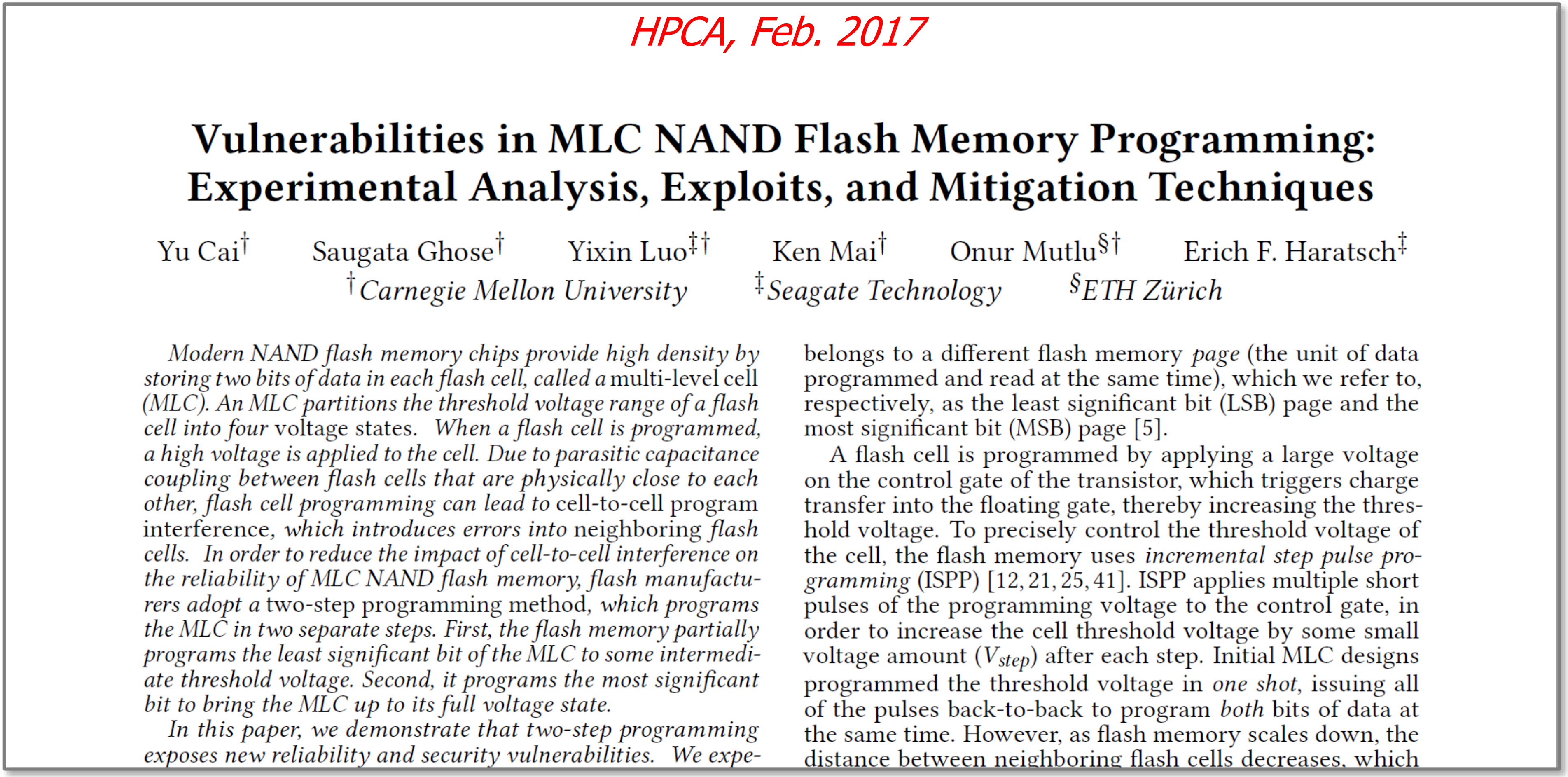
Justin Meza
Carnegie Mellon University
meza@cmu.edu

Qiang Wu
Facebook, Inc.
qwu@fb.com

Sanjeev Kumar
Facebook, Inc.
skumar@fb.com

Onur Mutlu
Carnegie Mellon University
onur@cmu.edu

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-

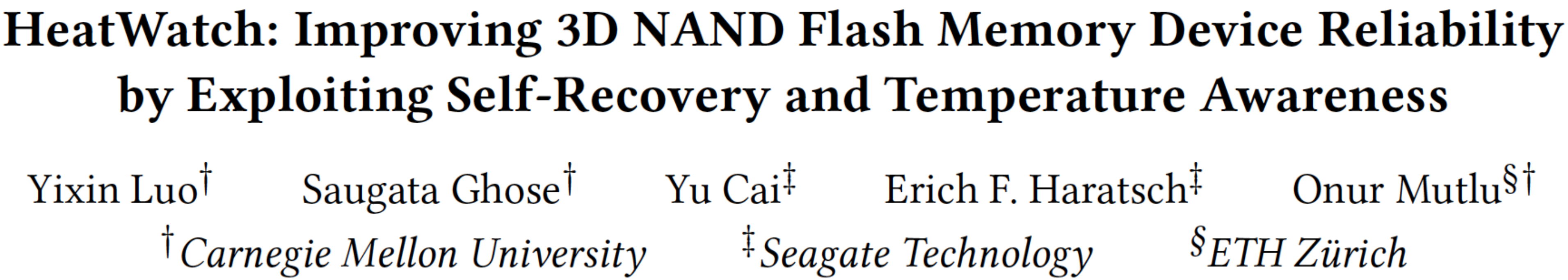
belongs to a different flash memory *page* (the unit of data programmed and read at the same time), which we refer to, respectively, as the least significant bit (LSB) page and the most significant bit (MSB) page [5].

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

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)]

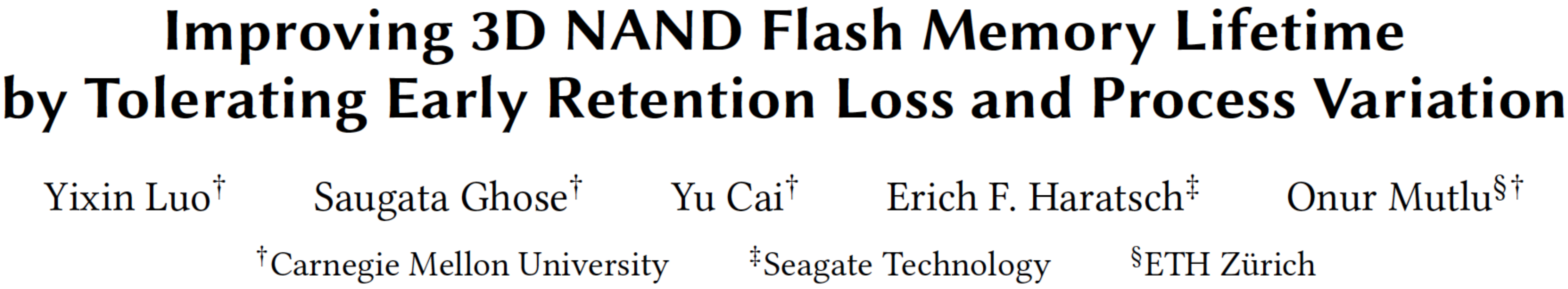


HeatWatch: Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature Awareness

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]
  [POMACS Journal Version (same content, different format)]
  [Slides (pptx) (pdf)]



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

Recall: Collapse of the “Galloping Gertie”

Source: AP
http://www.wsdot.wa.gov/tnbhistory/connections/connections3.htm

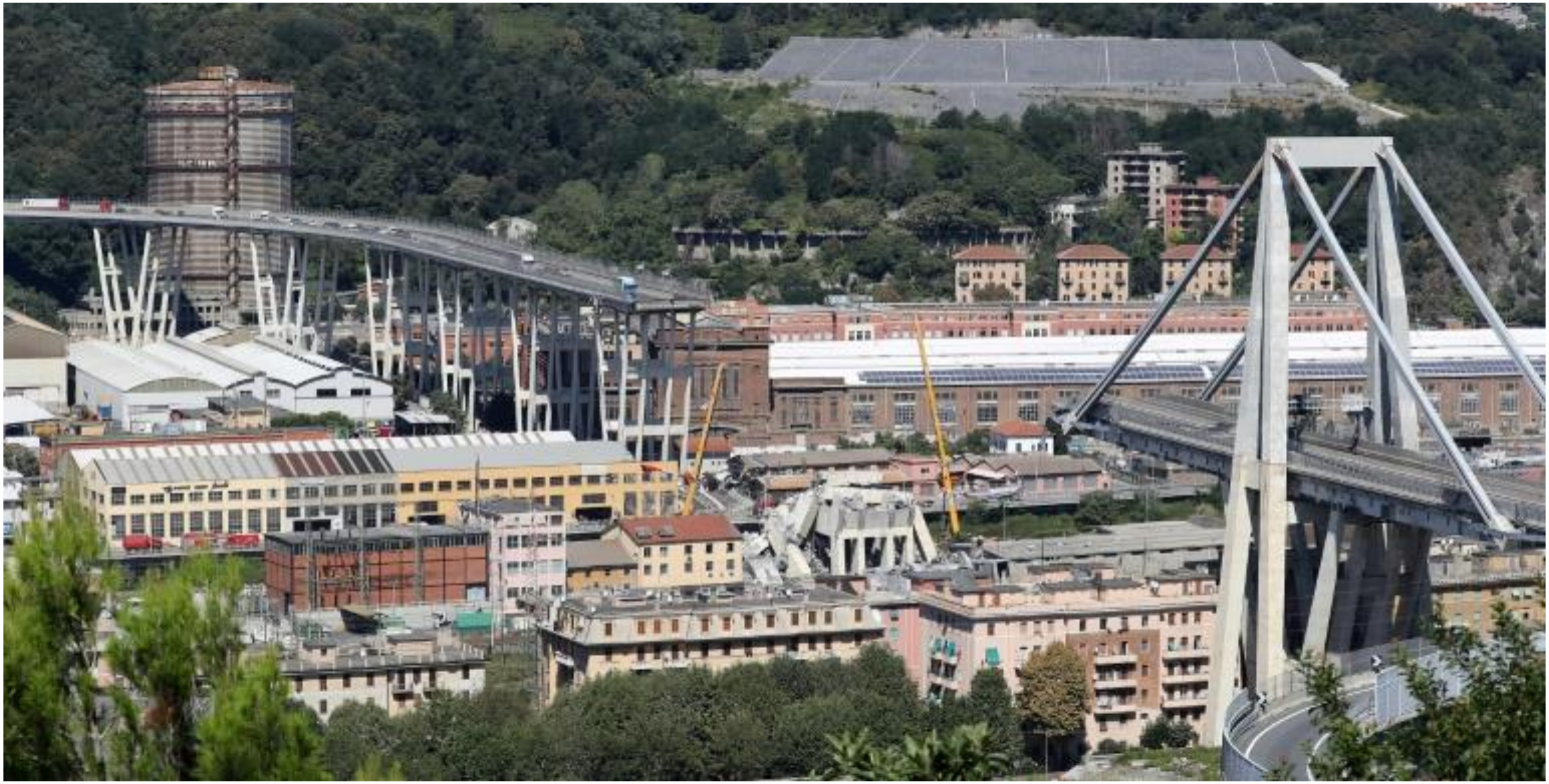
Another Example (1994)

Source: By 최광모 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=35197984

Yet Another Example (2007)

Source: Morry Gash/AP,
https://www.npr.org/2017/08/01/540669701/10-years-after-bridge-collapse-america-is-still-crumbling?t=1535427165809

A More Recent Example (2018)



Source: AFP / Valery HACHE, https://www.capitalfm.co.ke/news/2018/08/genoa-bridge-collapse-what-we-know/

The Takeaway, Again

In-Field Patch-ability

(Intelligent Memory)

Can Avoid Such Failures

Final Thoughts on RowHammer

Aside: Byzantine Failures

- This class of failures is known as Byzantine failures

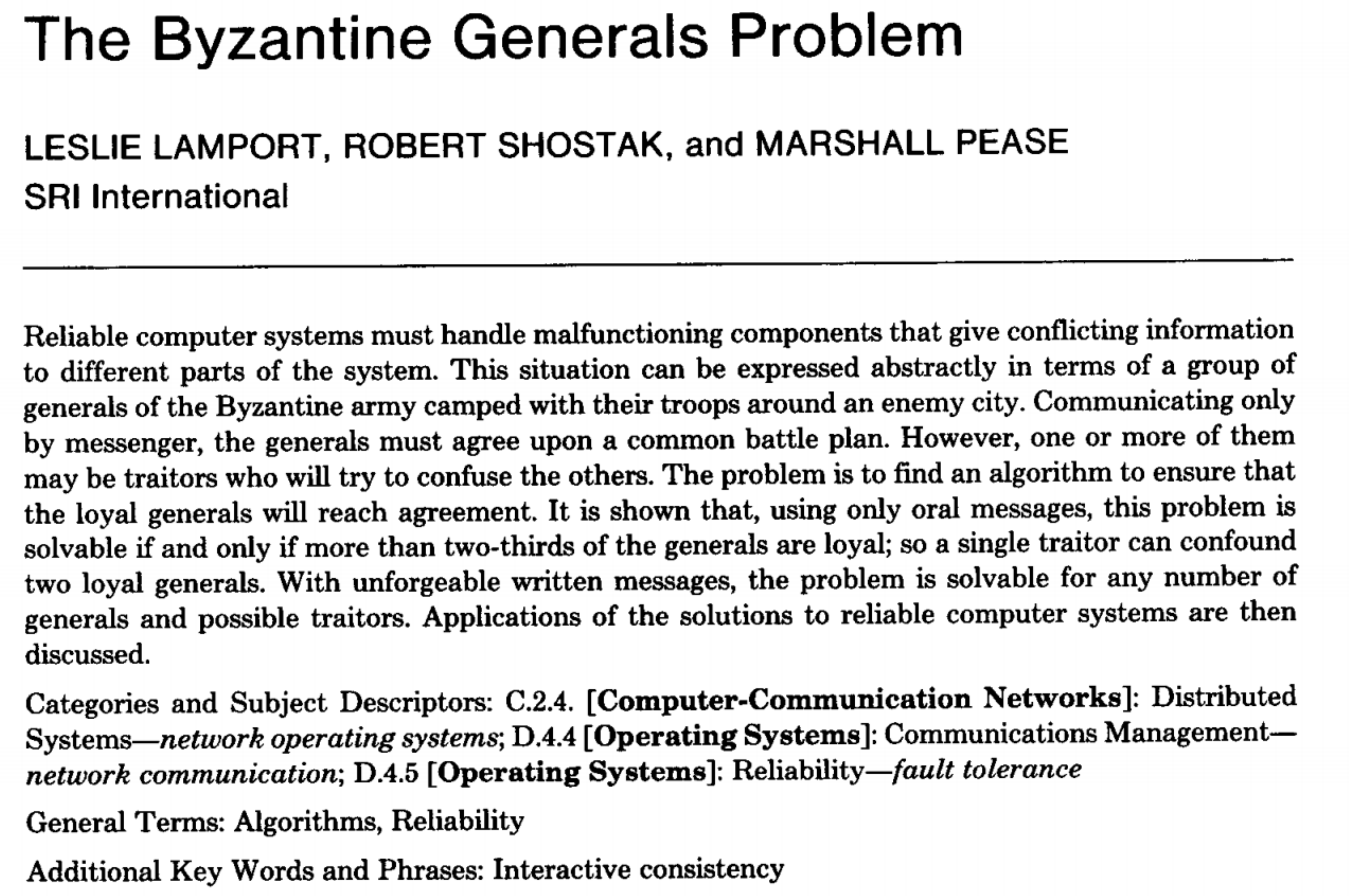
- Characterized by
  - Undetected erroneous computation
  - Opposite of “fail fast (with an error or no result)”

- “erroneous” can be “malicious” (intent is the only distinction)
- Very difficult to detect and confine Byzantine failures
- Do all you can to avoid them

- Lamport et al., “The Byzantine Generals Problem,” ACM TOPLAS 1982.

Slide credit: Mahadev Satyanarayanan, CMU, 15-440, Spring 2015

Aside: Byzantine Generals Problem



The Byzantine Generals Problem

LESLIE LAMPORT, ROBERT SHOSTAK, and MARSHALL PEASE
SRI International

Reliable computer systems must handle malfunctioning components that give conflicting information to different parts of the system. This situation can be expressed abstractly in terms of a group of generals of the Byzantine army camped with their troops around an enemy city. Communicating only by messenger, the generals must agree upon a common battle plan. However, one or more of them may be traitors who will try to confuse the others. The problem is to find an algorithm to ensure that the loyal generals will reach agreement. It is shown that, using only oral messages, this problem is solvable if and only if more than two-thirds of the generals are loyal; so a single traitor can confound two loyal generals. With unforgeable written messages, the problem is solvable for any number of generals and possible traitors. Applications of the solutions to reliable computer systems are then discussed.

Categories and Subject Descriptors: C.2.4. [**Computer-Communication Networks**]: Distributed Systems—*network operating systems*; D.4.4 [**Operating Systems**]: Communications Management—*network communication*; D.4.5 [**Operating Systems**]: Reliability—*fault tolerance*

General Terms: Algorithms, Reliability

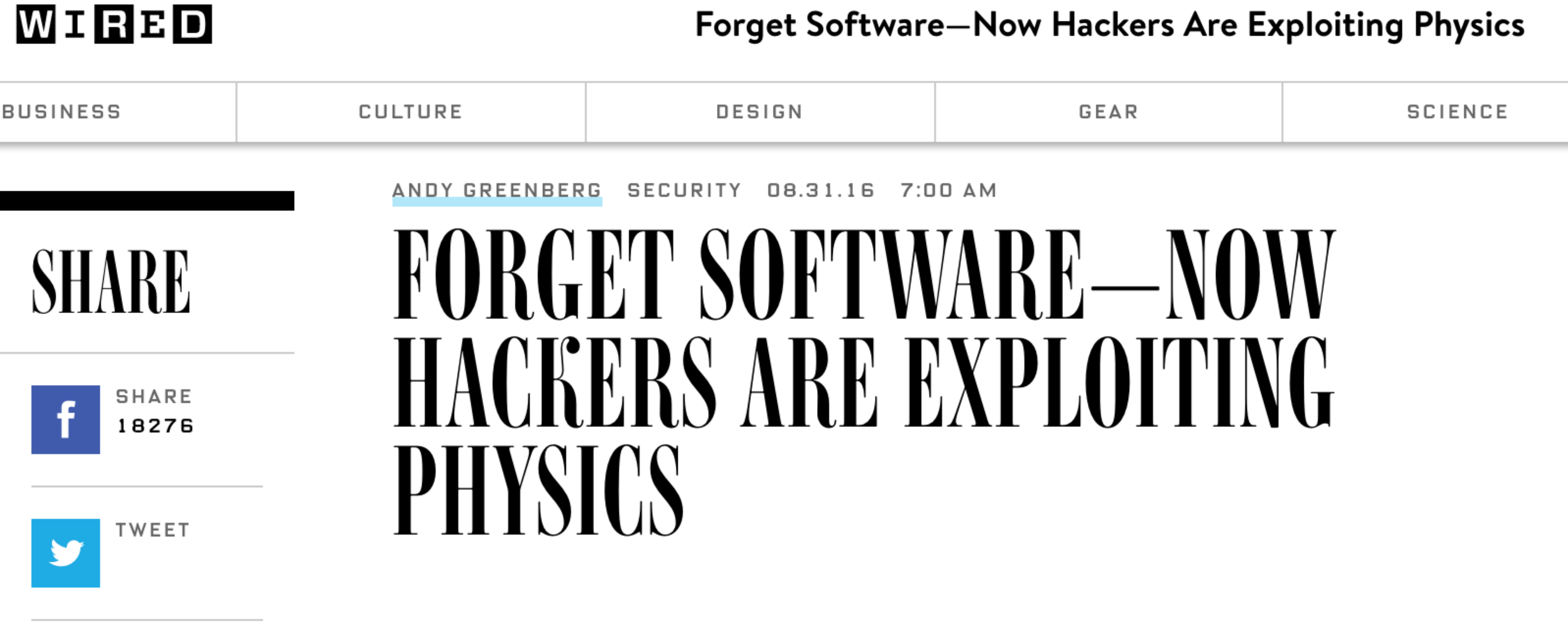
Additional Key Words and Phrases: Interactive consistency

https://dl.acm.org/citation.cfm?id=357176

RowHammer, Revisited

- One can predictably induce bit flips in commodity DRAM chips
  - >80% of the tested DRAM chips are vulnerable

- First example of how a simple hardware failure mechanism can create a widespread system security vulnerability



WIRED

Forget Software—Now Hackers Are Exploiting Physics

BUSINESS | CULTURE | DESIGN | GEAR | SCIENCE

ANDY GREENBERG SECURITY 08.31.16 7:00 AM

FORGET SOFTWARE—NOW HACKERS ARE EXPLOITING PHYSICS

RowHammer: Retrospective

- New mindset that has enabled a renewed interest in HW security attack research:
  - Real (memory) chips are vulnerable, in a simple and widespread manner → this causes real security problems
  - Hardware reliability → security connection is now mainstream discourse

- Many new RowHammer attacks…
  - Tens of papers in top security venues
  - **More to come** as RowHammer is getting worse (DDR4 & beyond)

- Many new RowHammer solutions…
  - Apple security release; Memtest86 updated
  - Many solution proposals in top venues (latest in ISCA 2019)
  - Principled system-DRAM co-design (in original RowHammer paper)
  - **More to come…**

Perhaps Most Importantly…

- RowHammer enabled a shift of mindset in mainstream security researchers
  - General-purpose hardware is fallible, in a widespread manner
  - Its problems are exploitable

- This mindset has enabled many systems security researchers to examine hardware in more depth
  - And understand HW's inner workings and vulnerabilities

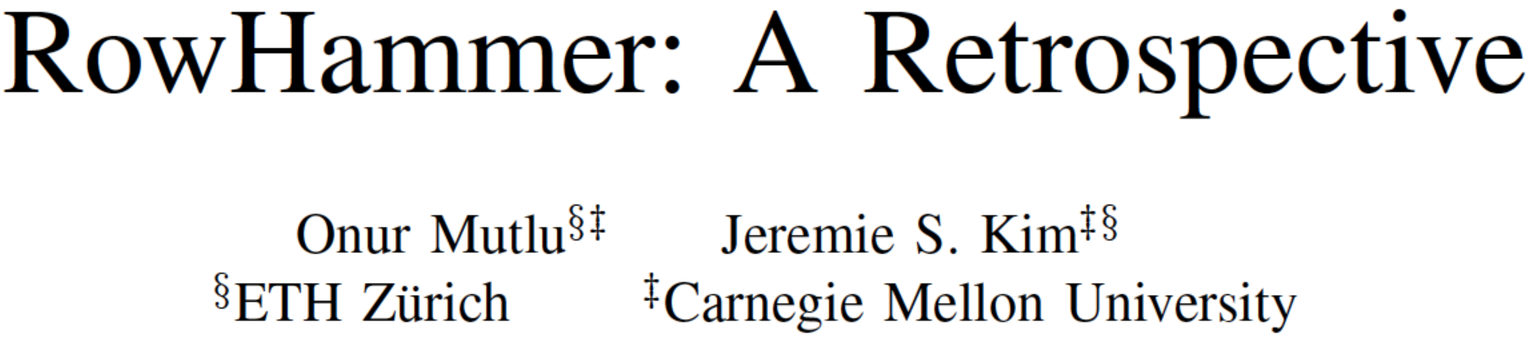
- It is no coincidence that two of the groups that discovered Meltdown and Spectre heavily worked on RowHammer attacks before
  - **More to come…**

Summary: RowHammer

- DRAM reliability is reducing
- Reliability issues open up security vulnerabilities
  - Very hard to defend against
- **Rowhammer is a prime example**
  - First example of how a simple hardware failure mechanism can create a widespread system security vulnerability
  - Its implications on system security research are tremendous & exciting
- Bad news: RowHammer is getting worse.
- **Good news: We have a lot more to do.**
  - We are now fully aware hardware is easily fallible.
  - We are developing both attacks and solutions.
  - We are developing principled models, methodologies, solutions.

For More on RowHammer…

- Onur Mutlu and Jeremie Kim,
  "RowHammer: A Retrospective"
  IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (***TCAD***) *Special Issue on Top Picks in Hardware and Embedded Security*, 2019.
  [Preliminary arXiv version]



RowHammer: A Retrospective

Onur Mutlu[§‡] Jeremie S. Kim[‡§]
[§]ETH Zürich [‡]Carnegie Mellon University

RowHammer in 2020 (I)

- Jeremie S. Kim, Minesh Patel, A. Giray Yaglikci, Hasan Hassan, Roknoddin Azizi, Lois Orosa, and Onur Mutlu,
  "Revisiting RowHammer: An Experimental Analysis of Modern Devices and Mitigation Techniques"
  Proceedings of the 47th International Symposium on Computer Architecture (***ISCA***), Valencia, Spain, June 2020.
  [Slides (pptx) (pdf)]
  [Lightning Talk Slides (pptx) (pdf)]
  [Talk Video (20 minutes)]
  [Lightning Talk Video (3 minutes)]

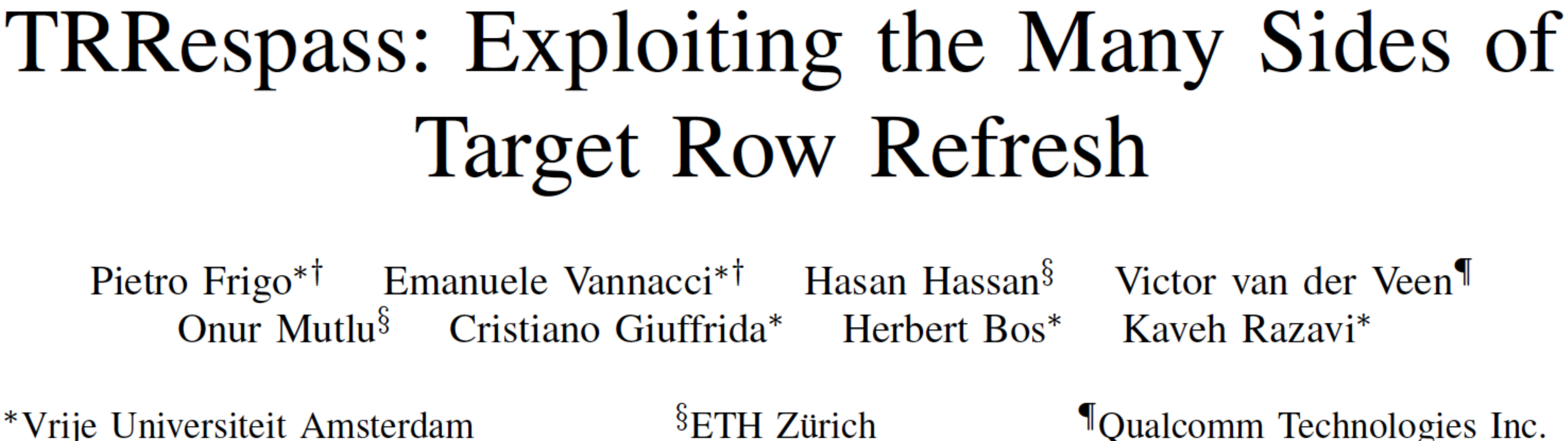
Revisiting RowHammer: An Experimental Analysis of Modern DRAM Devices and Mitigation Techniques

Jeremie S. Kim[§†] Minesh Patel[§] A. Giray Yağlıkçı[§]
Hasan Hassan[§] Roknoddin Azizi[§] Lois Orosa[§] Onur Mutlu[§†]

[§]*ETH Zürich* [†]*Carnegie Mellon University*

RowHammer in 2020 (II)

- Pietro Frigo, Emanuele Vannacci, Hasan Hassan, Victor van der Veen, Onur Mutlu, Cristiano Giuffrida, Herbert Bos, and Kaveh Razavi, **"TRRespass: Exploiting the Many Sides of Target Row Refresh"** *Proceedings of the 41st IEEE Symposium on Security and Privacy (**S&P**)*, San Francisco, CA, USA, May 2020.
  [Slides (pptx) (pdf)]
  [Talk Video (17 minutes)]
  [Source Code]
  [Web Article]
  Best paper award.



TRRespass: Exploiting the Many Sides of Target Row Refresh

Pietro Frigo*† Emanuele Vannacci*† Hasan Hassan§ Victor van der Veen¶
Onur Mutlu§ Cristiano Giuffrida* Herbert Bos* Kaveh Razavi*

*Vrije Universiteit Amsterdam §ETH Zürich ¶Qualcomm Technologies Inc.

RowHammer in 2020 (III)

- Lucian Cojocar, Jeremie Kim, Minesh Patel, Lillian Tsai, Stefan Saroiu, Alec Wolman, and Onur Mutlu,
  "Are We Susceptible to Rowhammer? An End-to-End Methodology for Cloud Providers"
  *Proceedings of the 41st IEEE Symposium on Security and Privacy (**S&P**)*, San Francisco, CA, USA, May 2020.
  [Slides (pptx) (pdf)]
  [Talk Video (17 minutes)]

Are We Susceptible to Rowhammer?
An End-to-End Methodology for Cloud Providers

Lucian Cojocar, Jeremie Kim[§†], Minesh Patel[§], Lillian Tsai[‡],
Stefan Saroiu, Alec Wolman, and Onur Mutlu[§†]
Microsoft Research, [§]ETH Zürich, [†]CMU, [‡]MIT

Rowhammer

Some History

Some More Historical Perspective

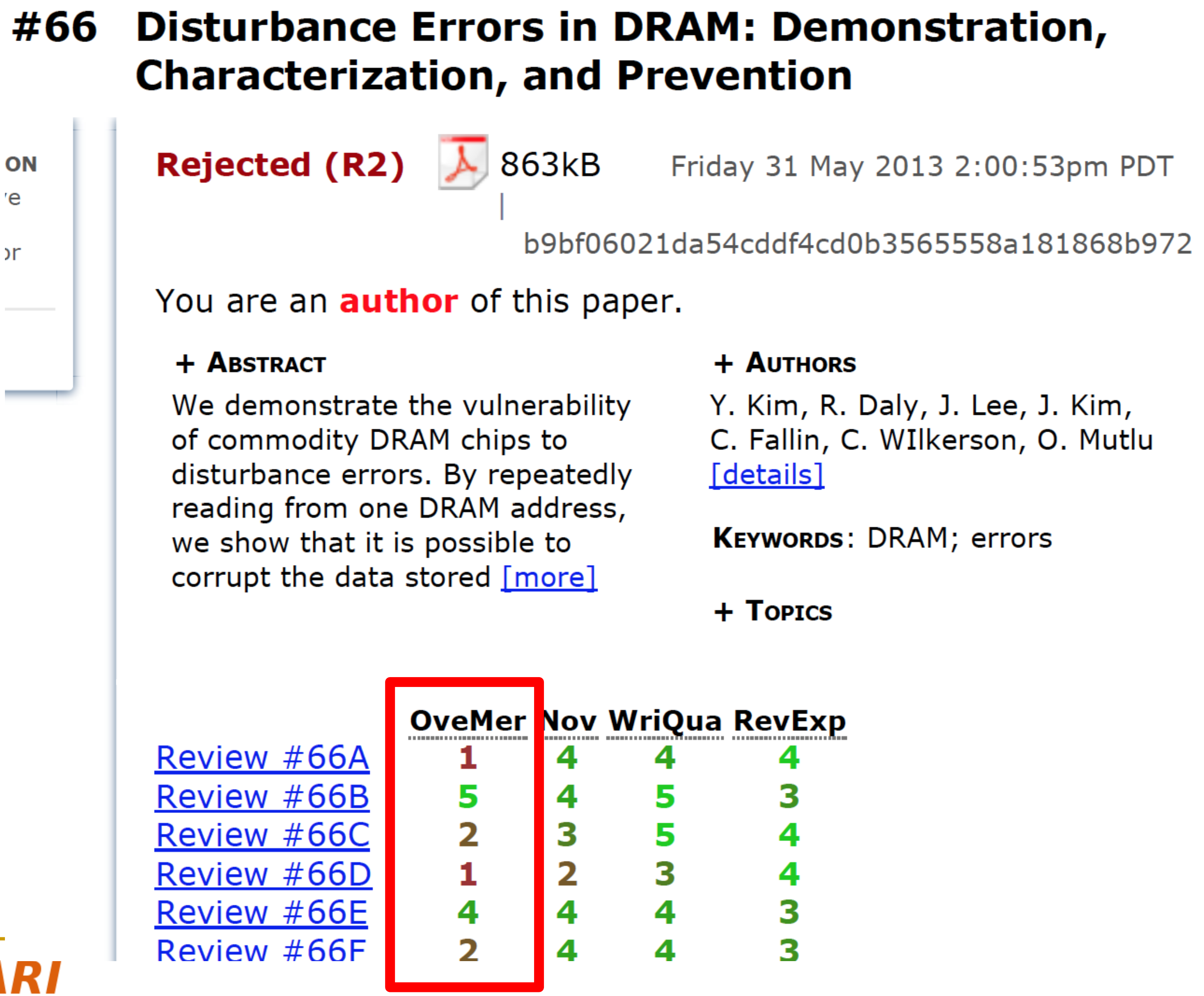
- RowHammer is the first example of a circuit-level failure mechanism causing a widespread system security vulnerability

- It led to a large body of work in security attacks, mitigations, architectural solutions, …

- Work building on RowHammer still continues
  - See MICRO 2020, ISCA 2020,

- Initially, it was dismissed by some reviewers
  - Rejected from MICRO 2013 conference

Initial RowHammer Reviews (MICRO 2013)



#66 Disturbance Errors in DRAM: Demonstration, Characterization, and Prevention

Rejected (R2) 863kB | Friday 31 May 2013 2:00:53pm PDT b9bf06021da54cddf4cd0b3565558a181868b972

You are an **author** of this paper.

+ Abstract

We demonstrate the vulnerability of commodity DRAM chips to disturbance errors. By repeatedly reading from one DRAM address, we show that it is possible to corrupt the data stored [more]

+ Authors

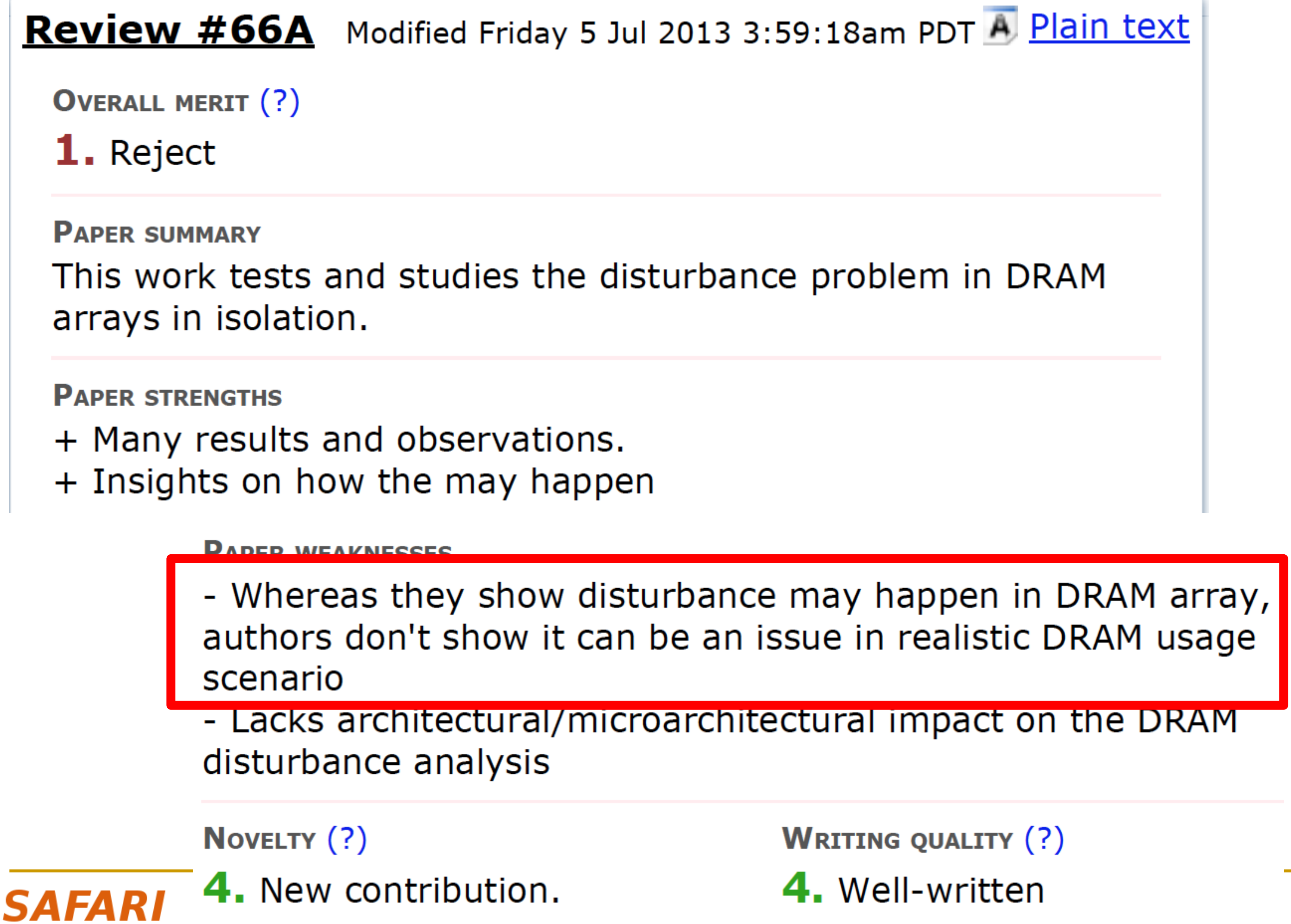
Y. Kim, R. Daly, J. Lee, J. Kim, C. Fallin, C. WIlkerson, O. Mutlu [details]

Keywords: DRAM; errors

+ Topics

	OveMer	Nov	WriQua	RevExp
Review #66A	1	4	4	4
Review #66B	5	4	5	3
Review #66C	2	3	5	4
Review #66D	1	2	3	4
Review #66E	4	4	4	3
Review #66F	2	4	4	3

Reviewer A



Review #66A Modified Friday 5 Jul 2013 3:59:18am PDT

OVERALL MERIT (?)

1. Reject

PAPER SUMMARY

This work tests and studies the disturbance problem in DRAM arrays in isolation.

PAPER STRENGTHS

+ Many results and observations.
+ Insights on how the may happen

PAPER WEAKNESSES

- Whereas they show disturbance may happen in DRAM array, authors don't show it can be an issue in realistic DRAM usage scenario
- Lacks architectural/microarchitectural impact on the DRAM disturbance analysis

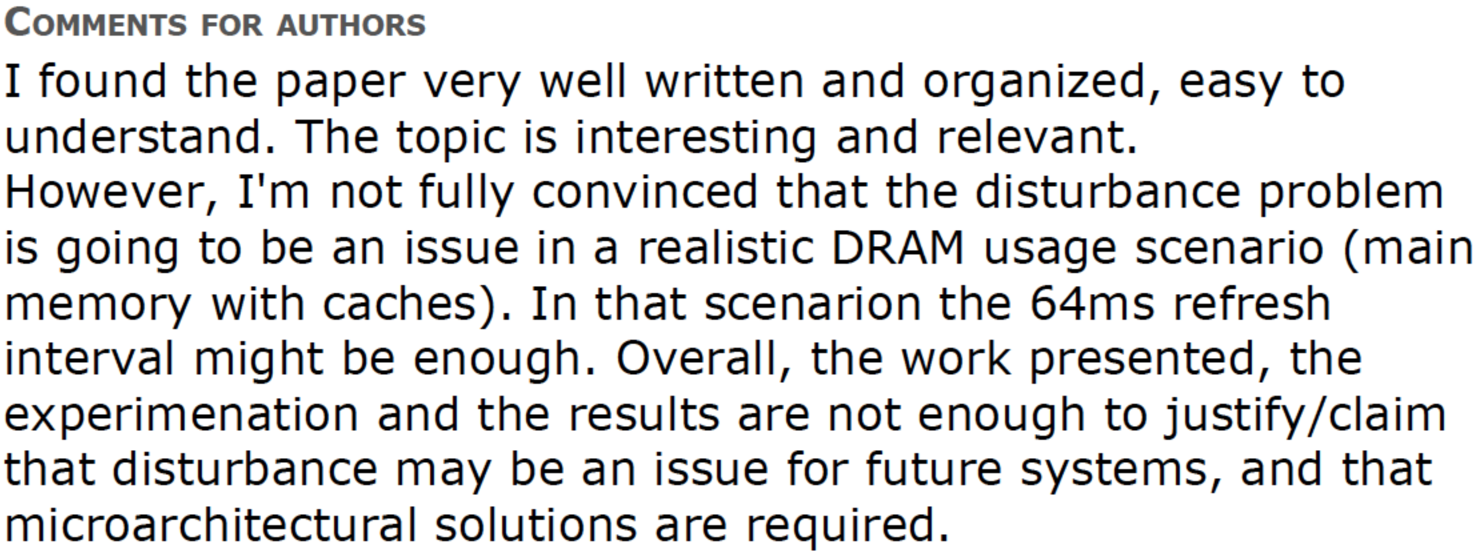
NOVELTY (?)

4. New contribution.

WRITING QUALITY (?)

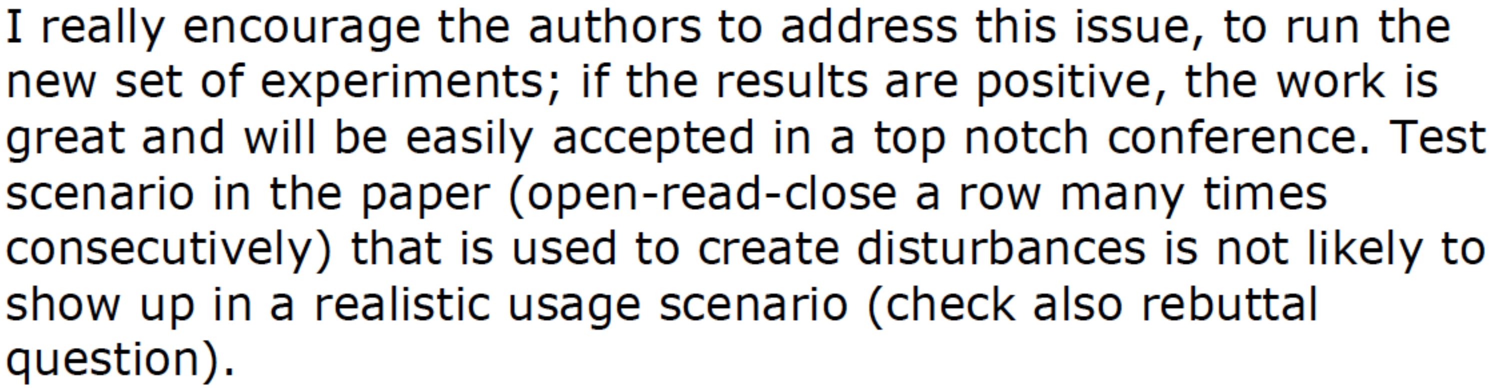
4. Well-written

Reviewer A -- Security is Not "Realistic"



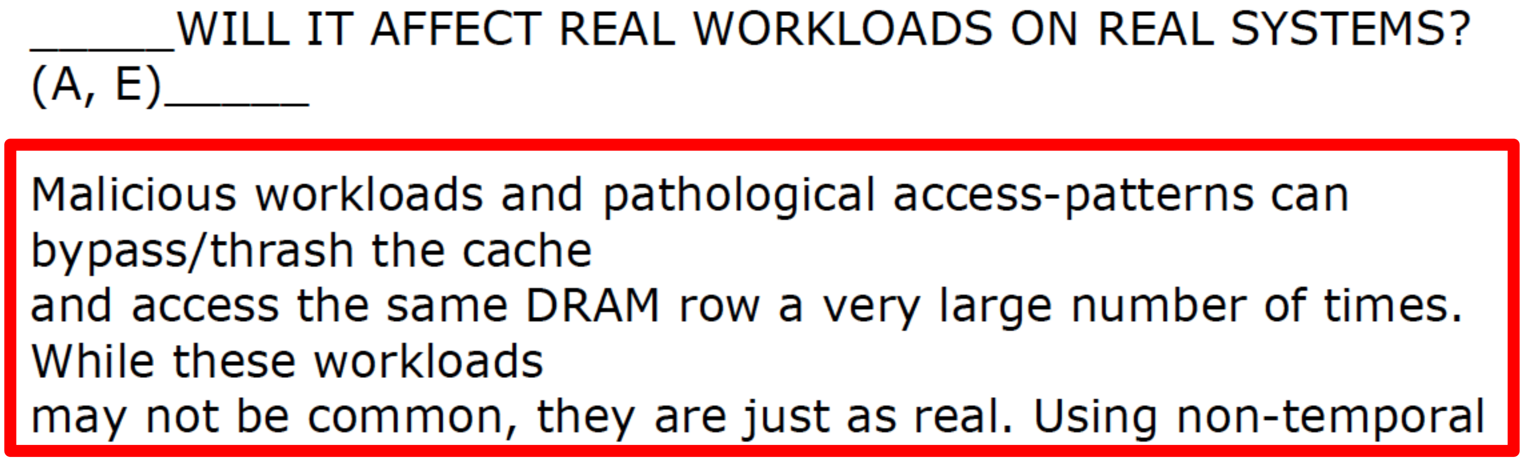
COMMENTS FOR AUTHORS

I found the paper very well written and organized, easy to understand. The topic is interesting and relevant.
However, I'm not fully convinced that the disturbance problem is going to be an issue in a realistic DRAM usage scenario (main memory with caches). In that scenarion the 64ms refresh interval might be enough. Overall, the work presented, the experimenation and the results are not enough to justify/claim that disturbance may be an issue for future systems, and that microarchitectural solutions are required.



I really encourage the authors to address this issue, to run the new set of experiments; if the results are positive, the work is great and will be easily accepted in a top notch conference. Test scenario in the paper (open-read-close a row many times consecutively) that is used to create disturbances is not likely to show up in a realistic usage scenario (check also rebuttal question).

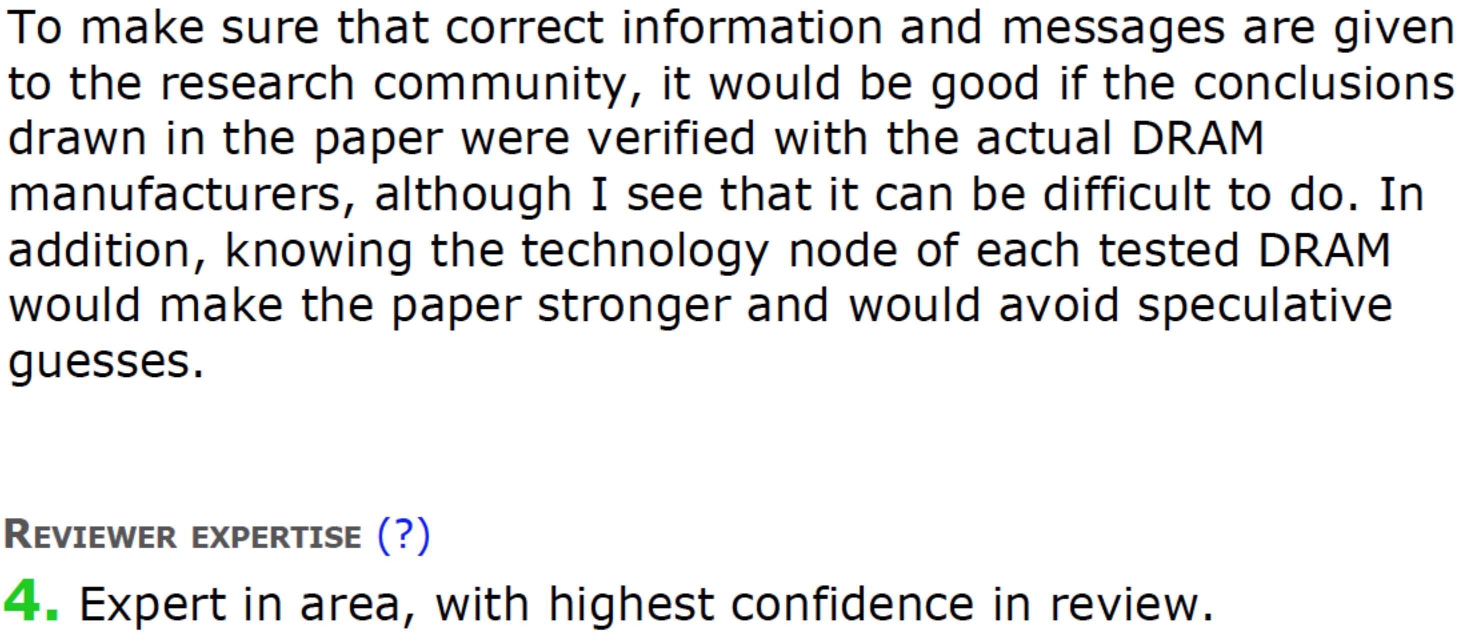
Rebuttal to Reviewer A



______WILL IT AFFECT REAL WORKLOADS ON REAL SYSTEMS? (A, E)______

Malicious workloads and pathological access-patterns can
bypass/thrash the cache
and access the same DRAM row a very large number of times.
While these workloads
may not be common, they are just as real. Using non-temporal

Reviewer A -- Demands

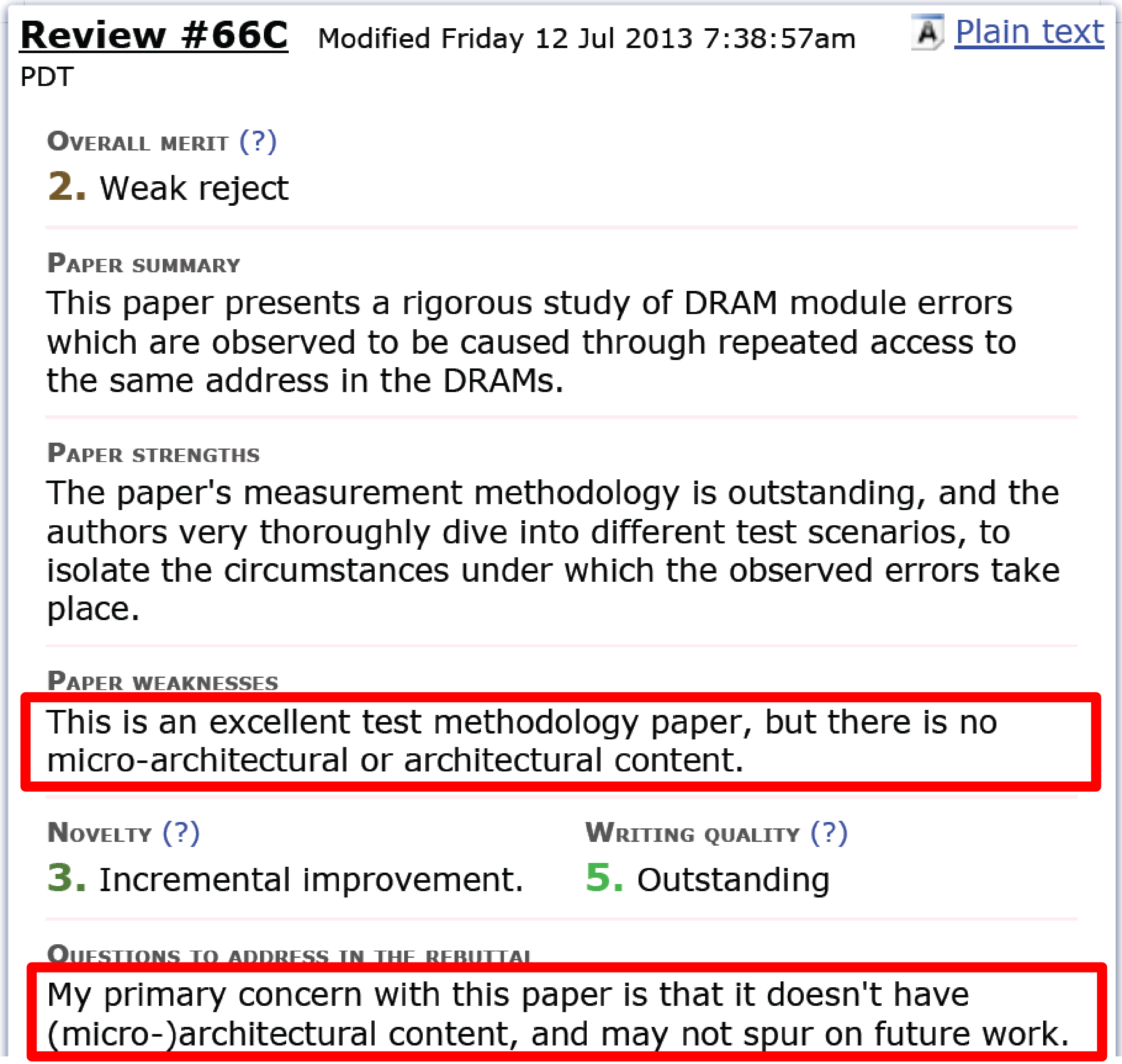


To make sure that correct information and messages are given to the research community, it would be good if the conclusions drawn in the paper were verified with the actual DRAM manufacturers, although I see that it can be difficult to do. In addition, knowing the technology node of each tested DRAM would make the paper stronger and would avoid speculative guesses.

REVIEWER EXPERTISE (?)

4. Expert in area, with highest confidence in review.

Reviewer C



Review #66C Modified Friday 12 Jul 2013 7:38:57am PDT

OVERALL MERIT (?)

2. Weak reject

PAPER SUMMARY

This paper presents a rigorous study of DRAM module errors which are observed to be caused through repeated access to the same address in the DRAMs.

PAPER STRENGTHS

The paper's measurement methodology is outstanding, and the authors very thoroughly dive into different test scenarios, to isolate the circumstances under which the observed errors take place.

PAPER WEAKNESSES

This is an excellent test methodology paper, but there is no micro-architectural or architectural content.

NOVELTY (?)

3. Incremental improvement.

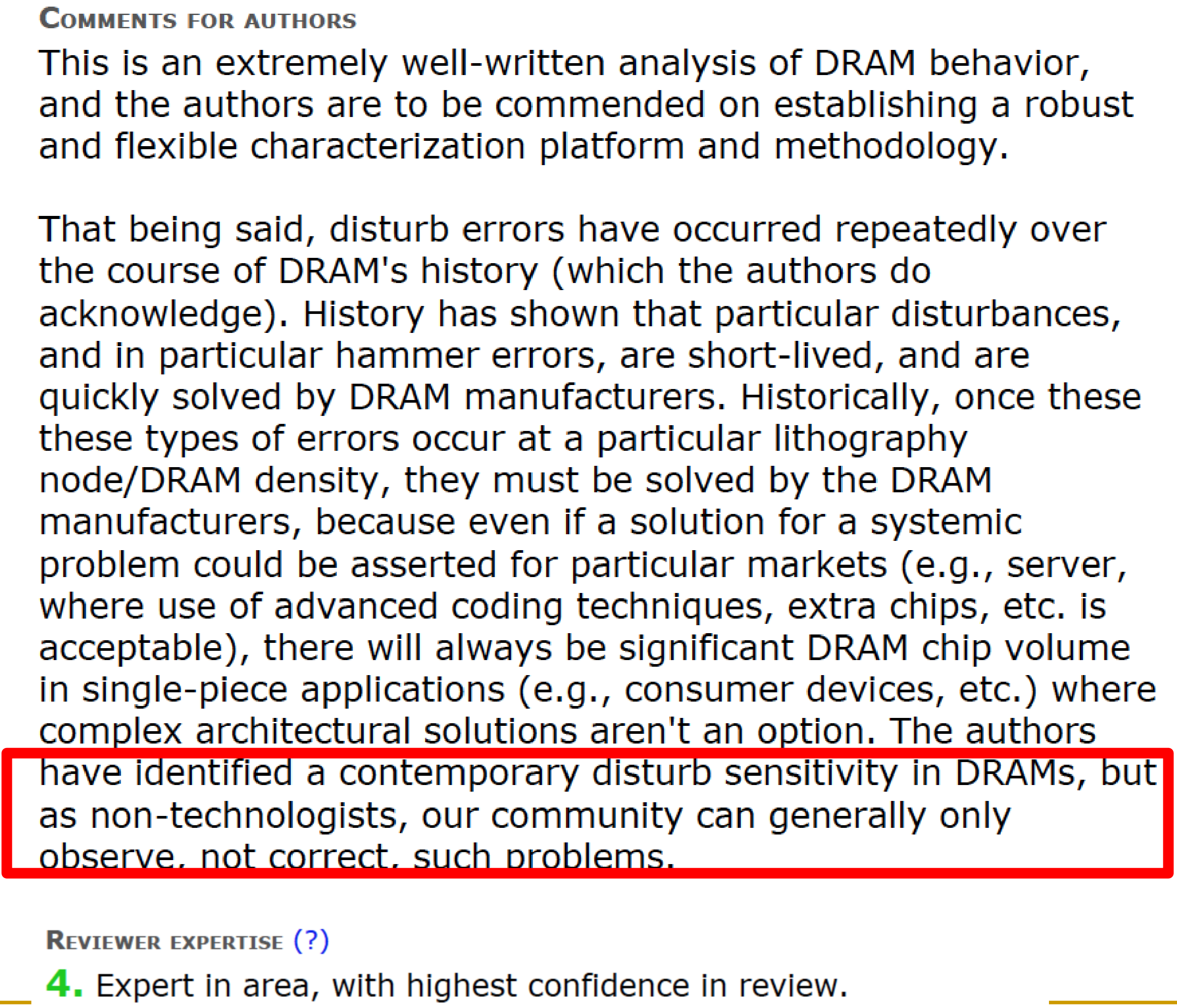
WRITING QUALITY (?)

5. Outstanding

QUESTIONS TO ADDRESS IN THE REBUTTAL

My primary concern with this paper is that it doesn't have (micro-)architectural content, and may not spur on future work.

Reviewer C -- Leave It to DRAM Vendors



Comments for authors

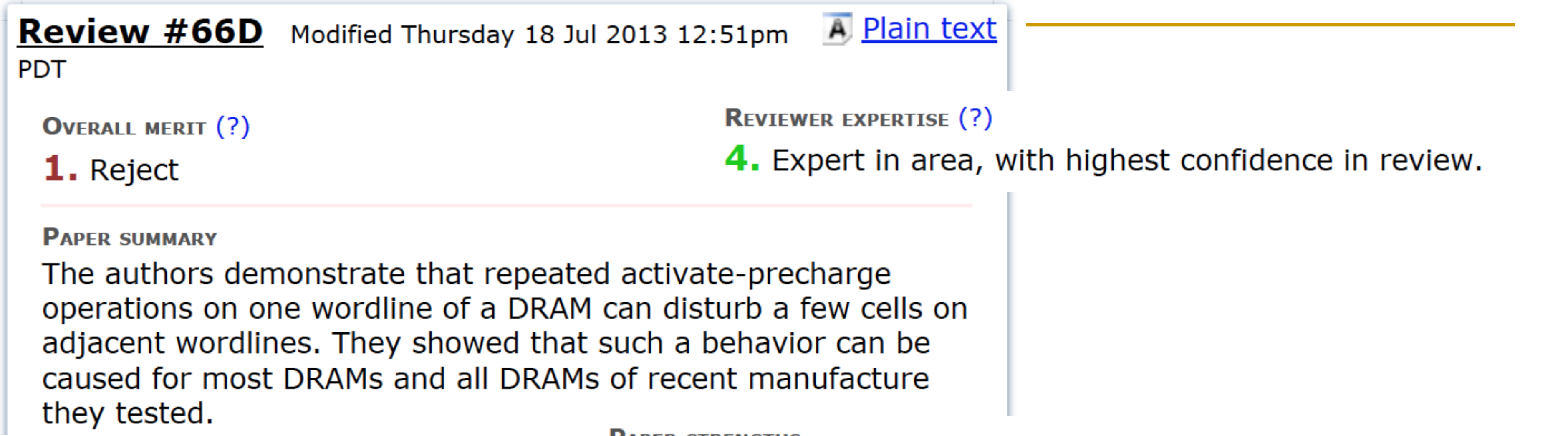
This is an extremely well-written analysis of DRAM behavior, and the authors are to be commended on establishing a robust and flexible characterization platform and methodology.

That being said, disturb errors have occurred repeatedly over the course of DRAM's history (which the authors do acknowledge). History has shown that particular disturbances, and in particular hammer errors, are short-lived, and are quickly solved by DRAM manufacturers. Historically, once these these types of errors occur at a particular lithography node/DRAM density, they must be solved by the DRAM manufacturers, because even if a solution for a systemic problem could be asserted for particular markets (e.g., server, where use of advanced coding techniques, extra chips, etc. is acceptable), there will always be significant DRAM chip volume in single-piece applications (e.g., consumer devices, etc.) where complex architectural solutions aren't an option. The authors have identified a contemporary disturb sensitivity in DRAMs, but as non-technologists, our community can generally only observe, not correct, such problems.

Reviewer expertise (?)

4. Expert in area, with highest confidence in review.

Reviewer D -- Nothing New in RowHammer



Review #66D Modified Thursday 18 Jul 2013 12:51pm PDT

OVERALL MERIT (?)

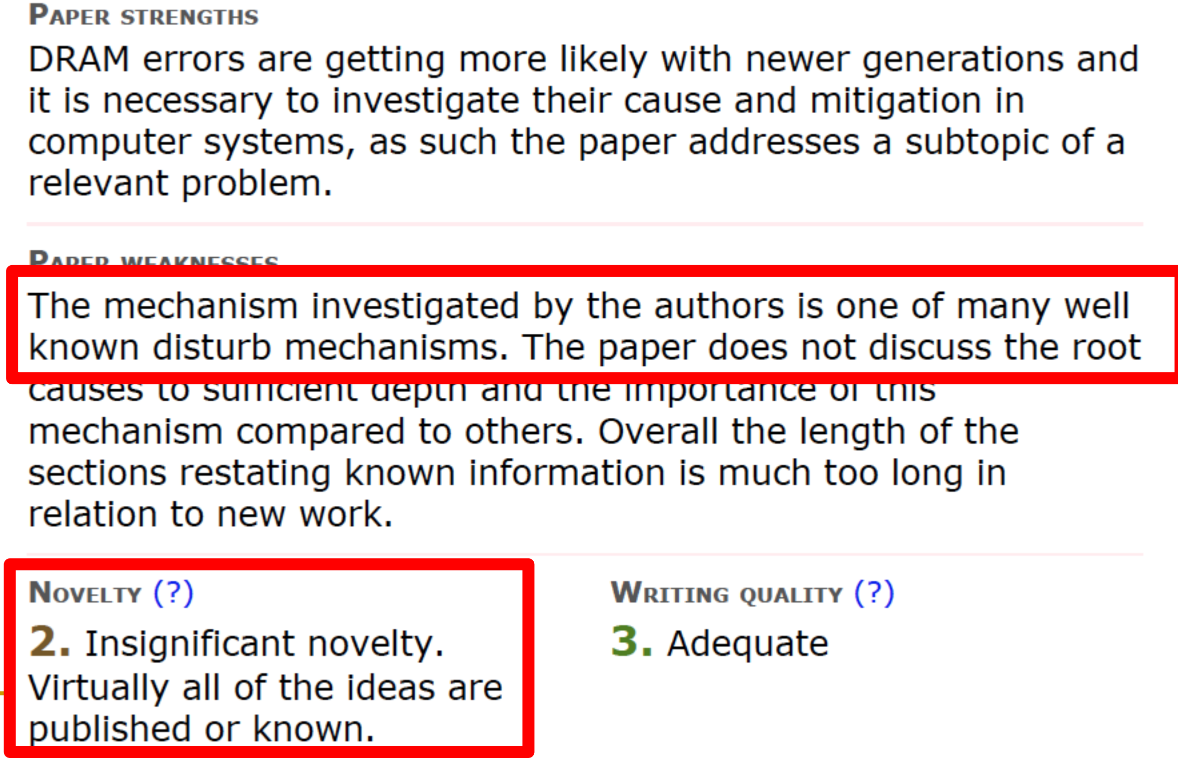
1. Reject

REVIEWER EXPERTISE (?)

4. Expert in area, with highest confidence in review.

PAPER SUMMARY

The authors demonstrate that repeated activate-precharge operations on one wordline of a DRAM can disturb a few cells on adjacent wordlines. They showed that such a behavior can be caused for most DRAMs and all DRAMs of recent manufacture they tested.



PAPER STRENGTHS

DRAM errors are getting more likely with newer generations and it is necessary to investigate their cause and mitigation in computer systems, as such the paper addresses a subtopic of a relevant problem.

PAPER WEAKNESSES

The mechanism investigated by the authors is one of many well known disturb mechanisms. The paper does not discuss the root causes to sufficient depth and the importance of this mechanism compared to others. Overall the length of the sections restating known information is much too long in relation to new work.

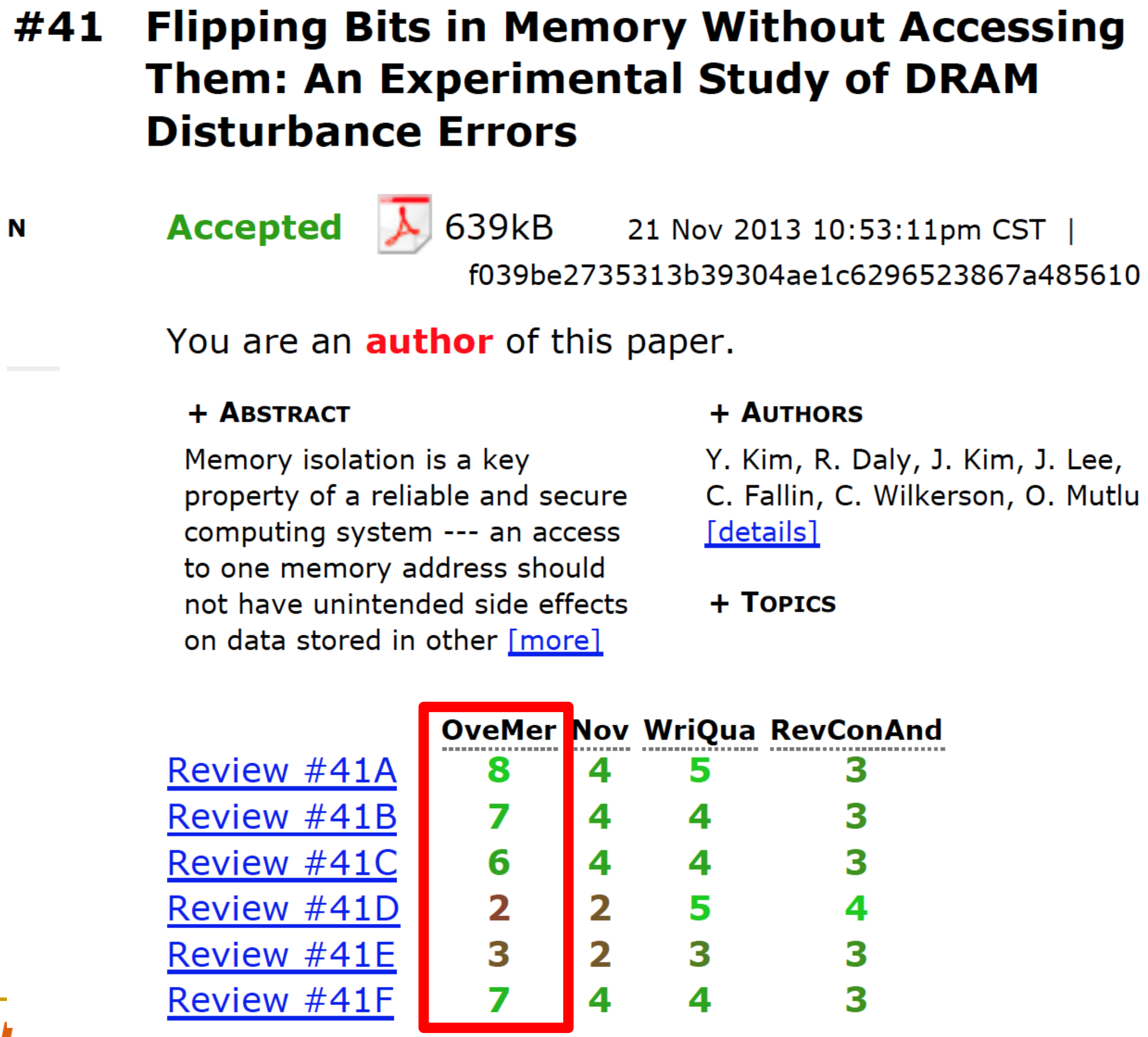
NOVELTY (?)

2. Insignificant novelty. Virtually all of the ideas are published or known.

WRITING QUALITY (?)

3. Adequate

ISCA 2014 Submission



#41 Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors

N

Accepted 639kB 21 Nov 2013 10:53:11pm CST | f039be2735313b39304ae1c6296523867a485610

You are an **author** of this paper.

+ Abstract

Memory isolation is a key property of a reliable and secure computing system --- an access to one memory address should not have unintended side effects on data stored in other [more]

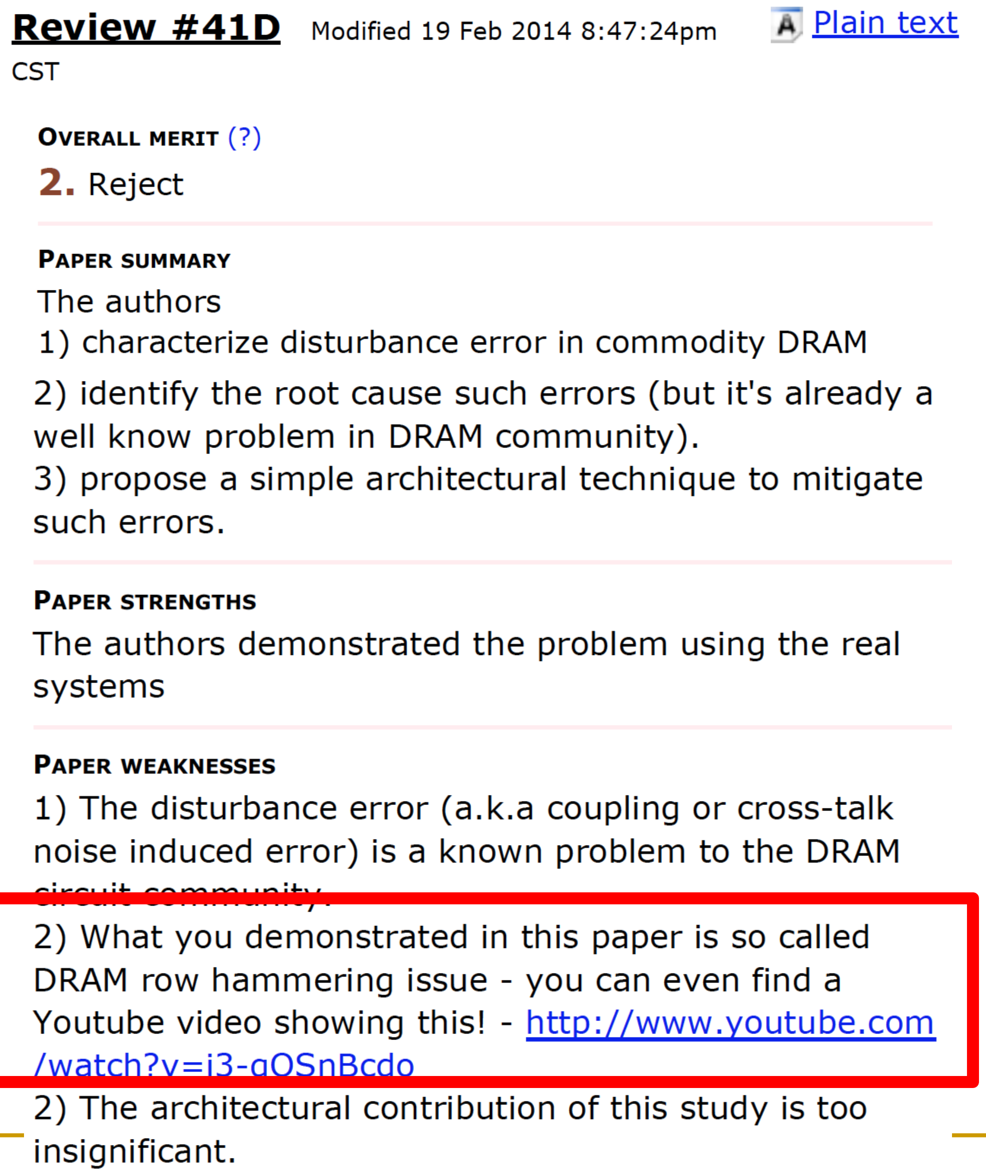
+ Authors

Y. Kim, R. Daly, J. Kim, J. Lee, C. Fallin, C. Wilkerson, O. Mutlu [details]

+ Topics

	OveMer	Nov	WriQua	RevConAnd
Review #41A	8	4	5	3
Review #41B	7	4	4	3
Review #41C	6	4	4	3
Review #41D	2	2	5	4
Review #41E	3	2	3	3
Review #41F	7	4	4	3

Reviewer D



Review #41D Modified 19 Feb 2014 8:47:24pm CST

Overall merit (?)

2. Reject

Paper summary

The authors

1) characterize disturbance error in commodity DRAM

2) identify the root cause such errors (but it's already a well know problem in DRAM community).

3) propose a simple architectural technique to mitigate such errors.

Paper strengths

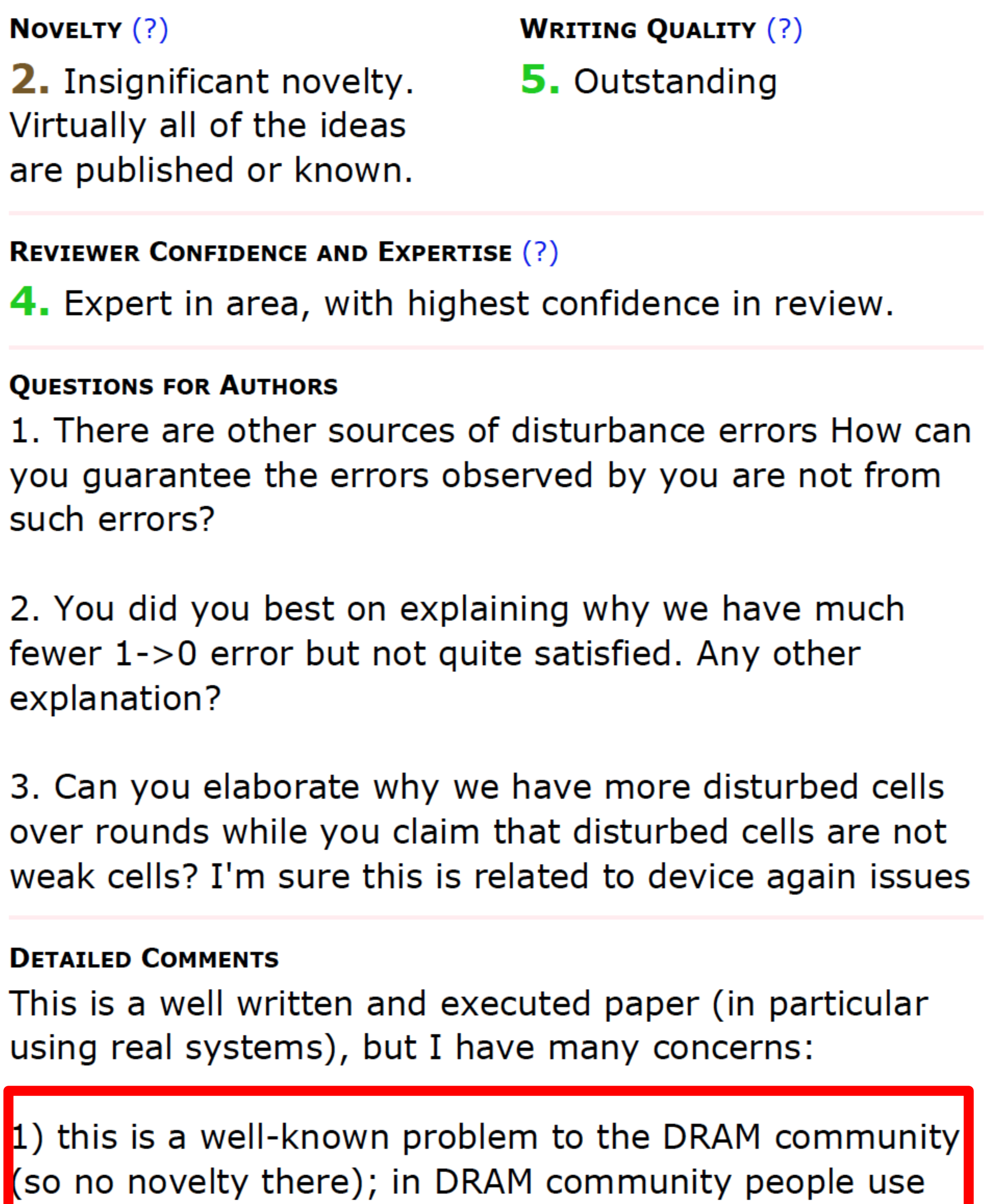
The authors demonstrated the problem using the real systems

Paper weaknesses

1) The disturbance error (a.k.a coupling or cross-talk noise induced error) is a known problem to the DRAM circuit community.

2) What you demonstrated in this paper is so called DRAM row hammering issue - you can even find a Youtube video showing this! - http://www.youtube.com/watch?v=i3-gQSnBcdo

2) The architectural contribution of this study is too insignificant.



Novelty (?)

2. Insignificant novelty. Virtually all of the ideas are published or known.

Writing Quality (?)

5. Outstanding

Reviewer Confidence and Expertise (?)

4. Expert in area, with highest confidence in review.

Questions for Authors

1. There are other sources of disturbance errors How can you guarantee the errors observed by you are not from such errors?

2. You did you best on explaining why we have much fewer 1->0 error but not quite satisfied. Any other explanation?

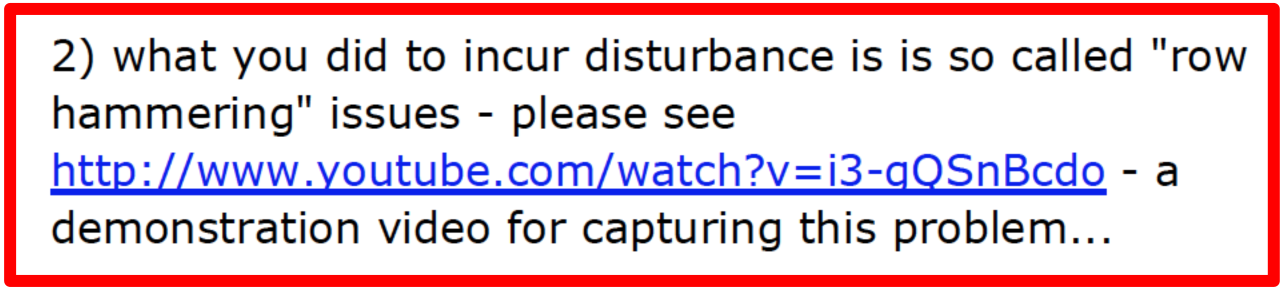
3. Can you elaborate why we have more disturbed cells over rounds while you claim that disturbed cells are not weak cells? I'm sure this is related to device again issues

Detailed Comments

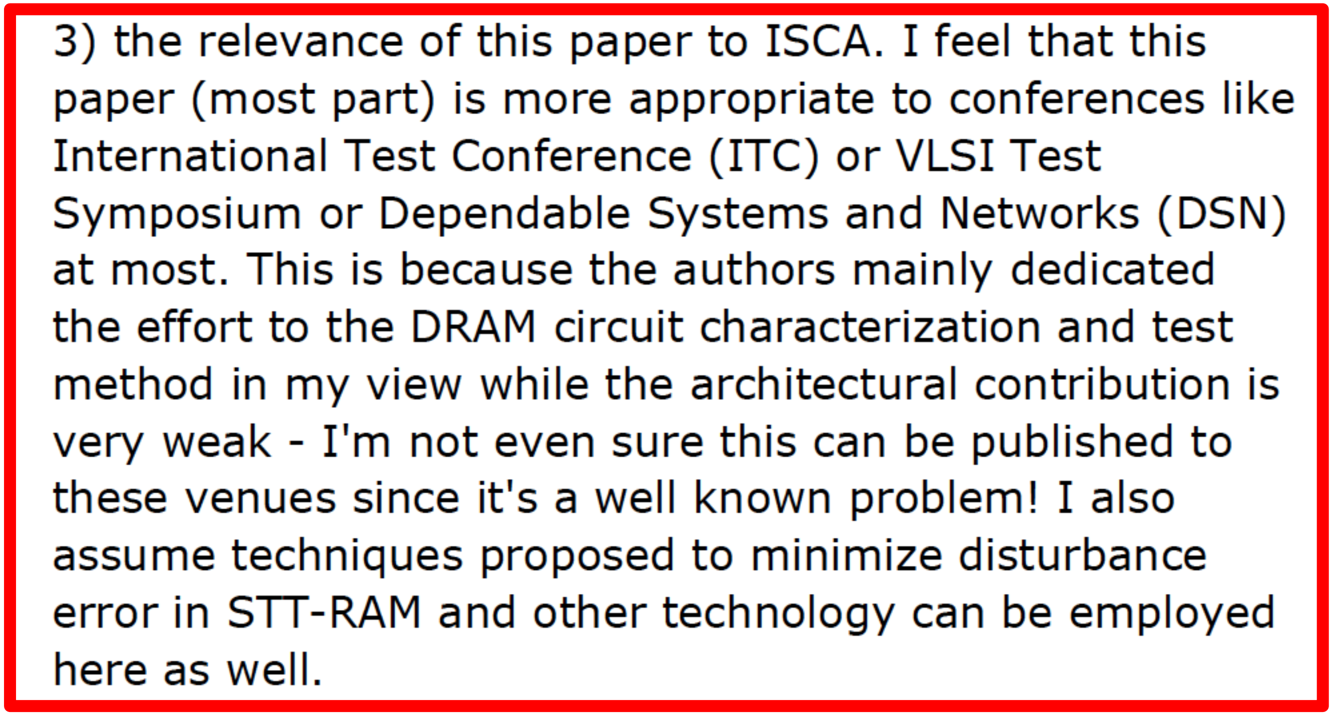
This is a well written and executed paper (in particular using real systems), but I have many concerns:

1) this is a well-known problem to the DRAM community (so no novelty there); in DRAM community people use

Reviewer D Continued…



2) what you did to incur disturbance is is so called "row hammering" issues - please see http://www.youtube.com/watch?v=i3-gQSnBcdo - a demonstration video for capturing this problem...



3) the relevance of this paper to ISCA. I feel that this paper (most part) is more appropriate to conferences like International Test Conference (ITC) or VLSI Test Symposium or Dependable Systems and Networks (DSN) at most. This is because the authors mainly dedicated the effort to the DRAM circuit characterization and test method in my view while the architectural contribution is very weak - I'm not even sure this can be published to these venues since it's a well known problem! I also assume techniques proposed to minimize disturbance error in STT-RAM and other technology can be employed here as well.

Rebuttal to Reviewer D

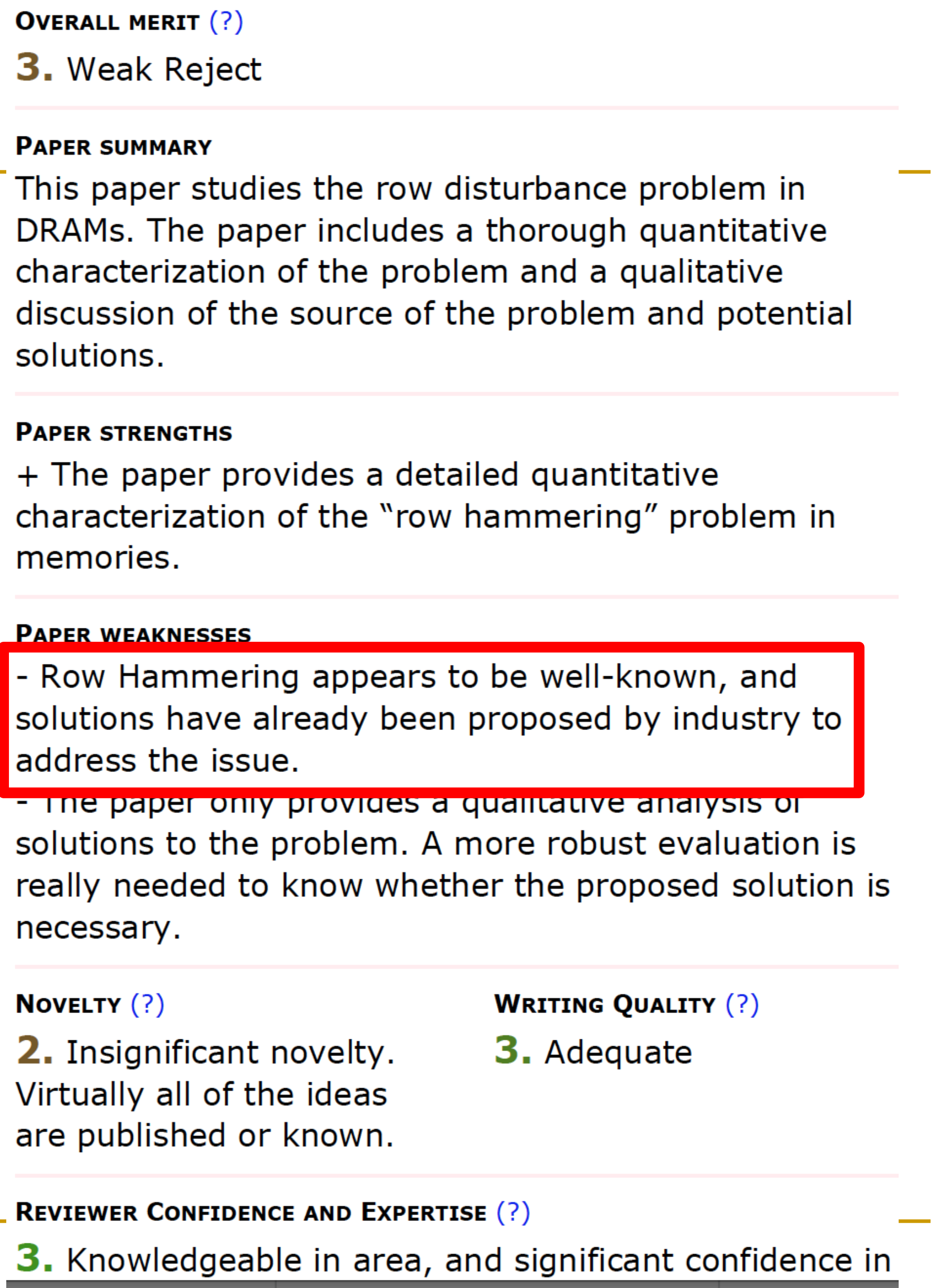
_____Reviewer D (Comments)_____

- 1. As we acknowledge in the paper, it is true that different types of DRAM coupling phenomena have been known to the DRAM circuits/testing community. However, there is a clear distinction between circuits/testing techniques confined to the *foundry* versus characterization/solution of a problem out in the *field*. The three citations (from 10+ years ago) do *not* demonstrate that disturbance errors exist in DIMMs sold then or now. They do *not* provide any real data (only simulated ones), let alone a large-scale characterization across many DIMMs from multiple manufacturers. They do *not* construct an attack on real systems, and they do *not* provide any solutions. Finally, our paper *already* references all three citations, or their more relevant equivalents. (The second/third citations provided by the reviewer are on bitline-coupling, whereas we cite works from the same authors on wordline-coupling [2, 3, 37].)

- 2. We were aware of the video from Teledyne (a test equipment company) and have *already* referenced slides from the same company [36]. In terms of their content regarding "row hammer", the video and the slides are identical: all they mention is that "aggressive row activations can corrupt adjacent rows". (They then advertise how their test equipment is able to capture a timestamped DRAM access trace, which can then be post-processed to identify when the number of activations exceeds a user-set threshold.) Both the video and slides do *not* say that this is a real problem affecting DIMMs on the market now. They do *not* provide any quantitative data, *nor* real-system demonstration, *nor* solution.

Reviewer E

Review #41E Modified 7 Feb 2014 11:08:04pm CST Plain text



Overall merit (?)

3. Weak Reject

Paper summary

This paper studies the row disturbance problem in DRAMs. The paper includes a thorough quantitative characterization of the problem and a qualitative discussion of the source of the problem and potential solutions.

Paper strengths

+ The paper provides a detailed quantitative characterization of the "row hammering" problem in memories.

Paper weaknesses

- Row Hammering appears to be well-known, and solutions have already been proposed by industry to address the issue.
- The paper only provides a qualitative analysis of solutions to the problem. A more robust evaluation is really needed to know whether the proposed solution is necessary.

Novelty (?)

2. Insignificant novelty. Virtually all of the ideas are published or known.

Writing Quality (?)

3. Adequate

Reviewer Confidence and Expertise (?)

3. Knowledgeable in area, and significant confidence in

but there are numerous mentions of hammering in the literature, and clearly industry has studied this problem for many years. In particular, Intel has a patent application on a memory controller technique that addresses this exact problem, with priority date June 2012:

http://www.google.com/patents/WO2014004748A1?cl=en

The patent application details sound very similar to solution 6 in this paper, so a more thorough comparison with solution 7 seems mandatory.

My overall feeling is that while the reliability characterization is important and interesting, a better target audience for the characterization work would be in a testing/reliability venue. The most interesting part of this paper from the ISCA point of view are the proposed solutions, but all of these are discussed in a very qualitative manner. My preference would be to see a much shorter characterization section with a much stronger and quantitative evaluation and comparison of the proposed solutions.

Rebuttal to Reviewer

_____Reviewer E (Comments)_____

After our paper was submitted, two patents that had been filed by Intel were made public (one is mentioned by the reviewer [R1]). Together, the two patents describe what we posed as the *sixth* potential solution in our paper (Section 8). Essentially, the memory controller maintains a table of counters to track the number of activations to recently activated rows [R2]. And if one of the counters exceeds a certain threshold, the memory controller notifies the DRAM chips using a special command [R1]. The DRAM chips would then refresh an entire "region" of rows that includes both the aggressor and its victim(s) [R1]. For the patent [R1] to work, DRAM manufacturers must cooperate and implement this special command. (It is a convenient way of circumventing the opacity in the logical-physical mapping. If implemented, the same command can also be used for our *seventh* solution.) The limitation of this *sixth* solution is the storage overhead of the counters and the extra power required to associatively search through them on every activation (Section 8). That is why we believe our *seventh* solution to be more attractive. We will cite the patents and include a more concrete comparison between the two solutions.

Nevertheless, we were able to induce a large number of DRAM disturbance errors on all the latest Intel/AMD platforms that we tested: Haswell, Ivy Bridge, Sandy Bridge, and Piledriver. (At the time of submission, we had tested only Sandy Bridge.) Importantly, the patents do *not* provide quantitative characterization

nor real-system demonstration.

[R1] "Row Hammer Refresh Command." US20140006703 A1
[R2] "Row Hammer Condition Monitoring." US20140006704 A1

Suggestions to Reviewers

- Be fair; you do not know it all
- Be open-minded; you do not know it all
- Be accepting of diverse research methods: there is no single way of doing research
- Be constructive, not destructive
- Do not have double standards…

Do not block or delay scientific progress for non-reasons

An Interview on Research and Education

- Computing Research and Education (@ ISCA 2019)
  - https://www.youtube.com/watch?v=8ffSEKZhmvo&list=PL5Q2soXY2Zi_4oP9LdL3cc8G6NIjD2Ydz

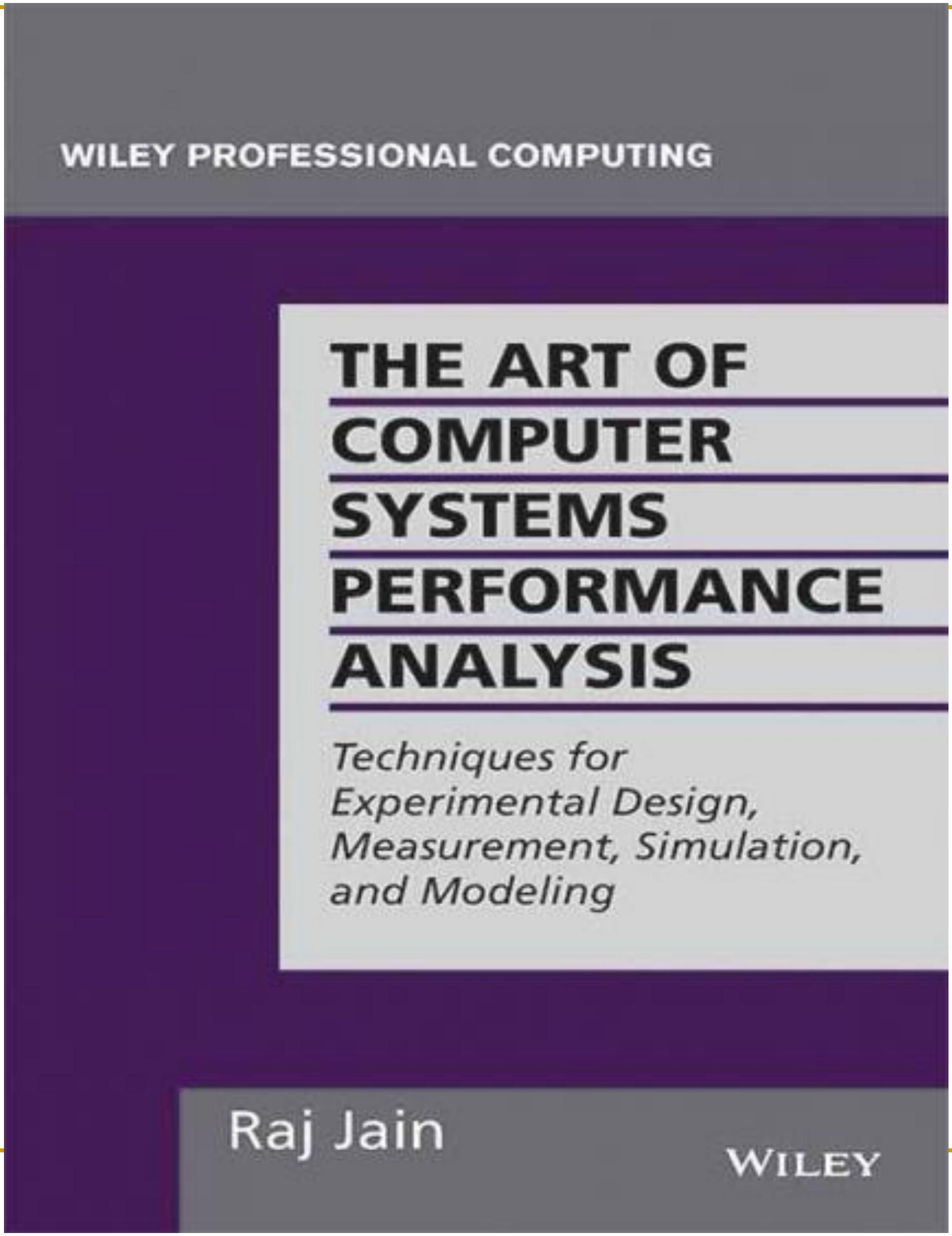
- Maurice Wilkes Award Speech (10 minutes)
  - https://www.youtube.com/watch?v=tcQ3zZ3JpuA&list=PL5Q2soXY2Zi8D_5MGV6EnXEJHnV2YFBJl&index=15

More Thoughts and Suggestions

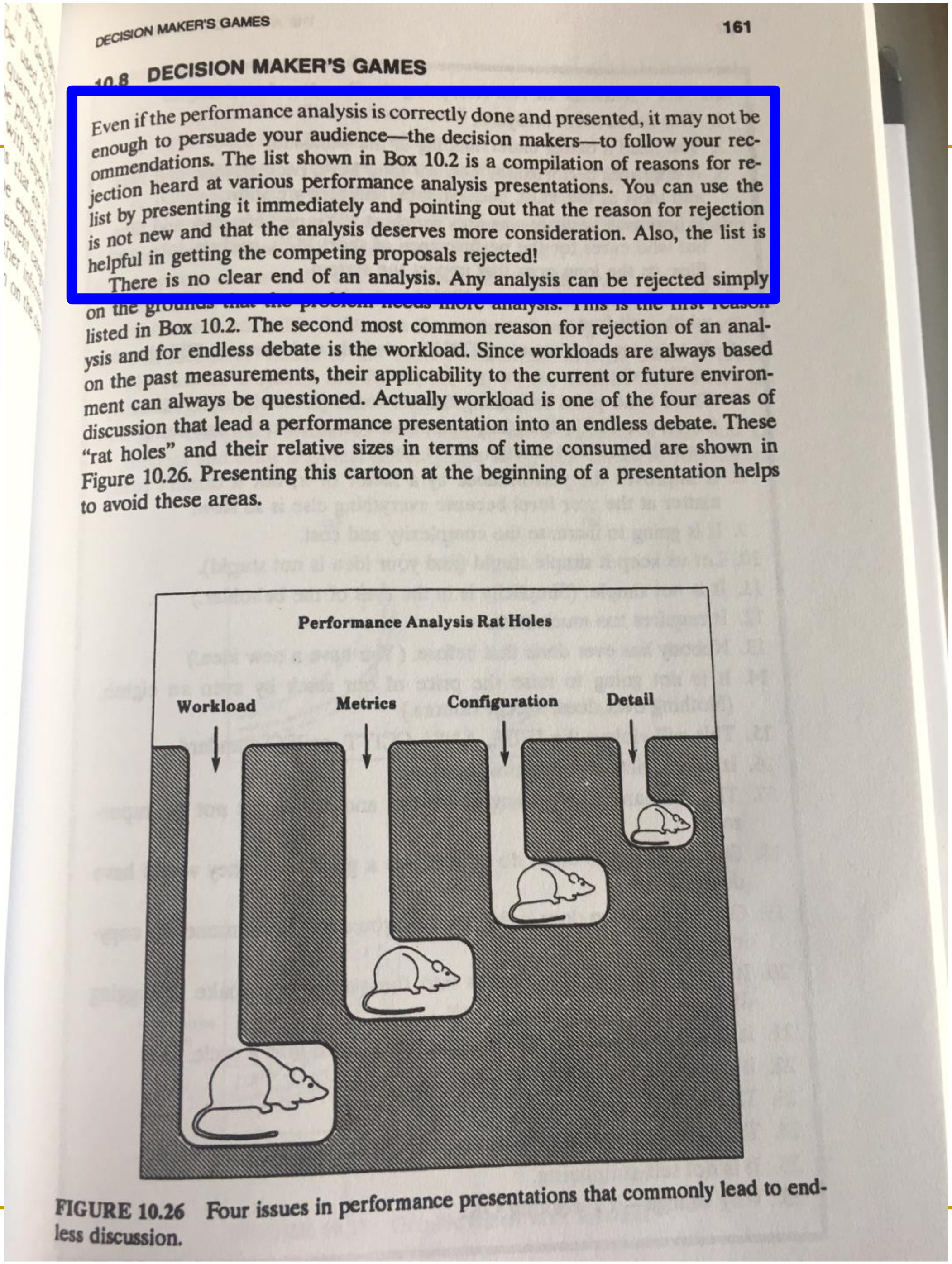
- Onur Mutlu,
  "Some Reflections (on DRAM)"
  *Award Speech for ACM SIGARCH Maurice Wilkes Award, at the **ISCA** Awards Ceremony*, Phoenix, AZ, USA, 25 June 2019.
  [Slides (pptx) (pdf)]
  [Video of Award Acceptance Speech (Youtube; 10 minutes) (Youku; 13 minutes)]
  [Video of Interview after Award Acceptance (Youtube; 1 hour 6 minutes) (Youku; 1 hour 6 minutes)]
  [News Article on "ACM SIGARCH Maurice Wilkes Award goes to Prof. Onur Mutlu"]

- Onur Mutlu,
  "How to Build an Impactful Research Group"
  *57th Design Automation Conference Early Career Workshop (**DAC**)*, Virtual, 19 July 2020.
  [Slides (pptx) (pdf)]

Aside: A Recommended Book


WILEY PROFESSIONAL COMPUTING
THE ART OF COMPUTER SYSTEMS PERFORMANCE ANALYSIS
Techniques for Experimental Design, Measurement, Simulation, and Modeling
Raj Jain
WILEY

Raj Jain, "The Art of Computer Systems Performance Analysis," Wiley, 1991.



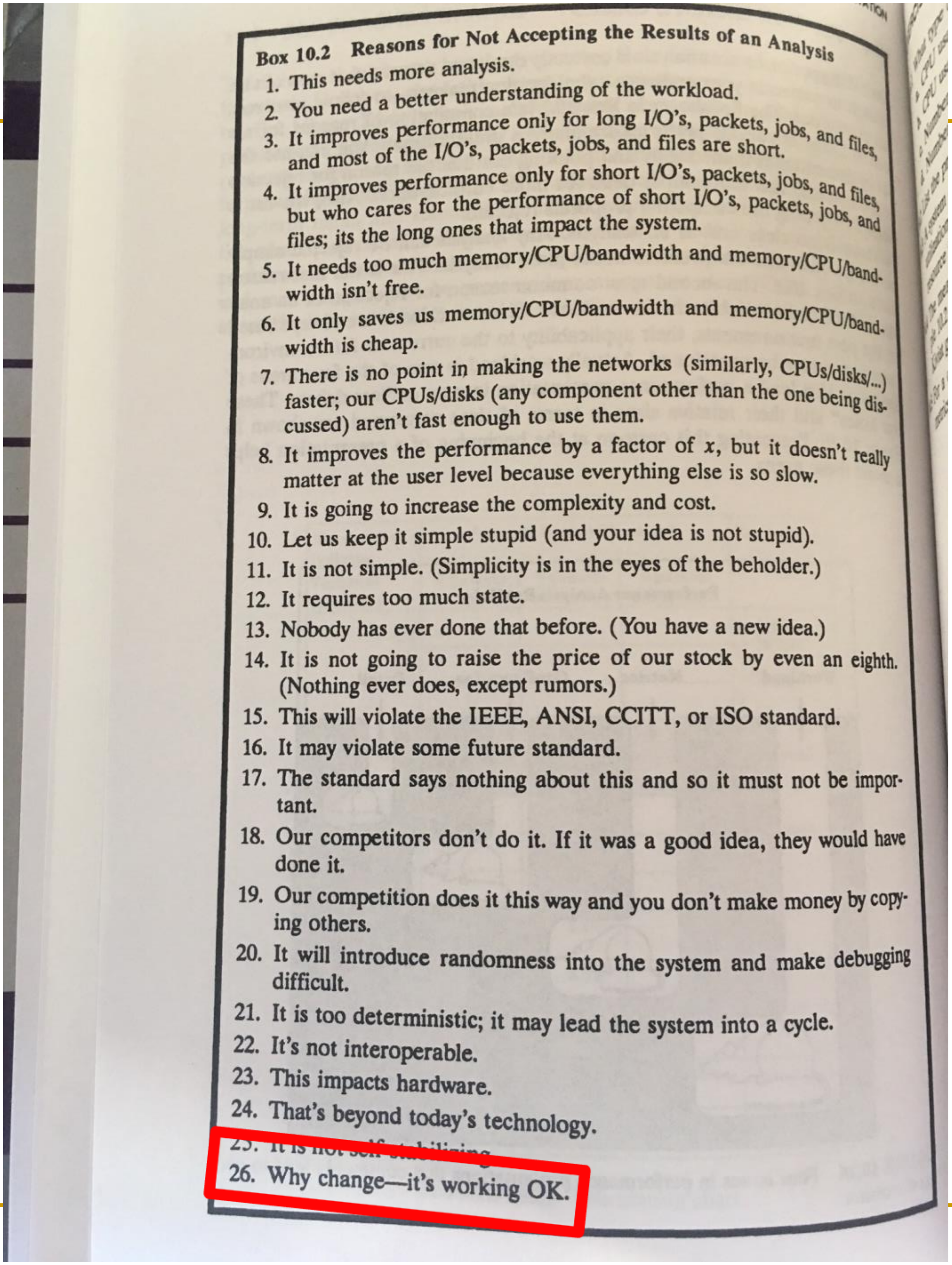
10.8 DECISION MAKER'S GAMES

Even if the performance analysis is correctly done and presented, it may not be enough to persuade your audience—the decision makers—to follow your recommendations. The list shown in Box 10.2 is a compilation of reasons for rejection heard at various performance analysis presentations. You can use the list by presenting it immediately and pointing out that the reason for rejection is not new and that the analysis deserves more consideration. Also, the list is helpful in getting the competing proposals rejected!

There is no clear end of an analysis. Any analysis can be rejected simply on the grounds that the problem needs more analysis. This is the first reason listed in Box 10.2. The second most common reason for rejection of an analysis and for endless debate is the workload. Since workloads are always based on the past measurements, their applicability to the current or future environment can always be questioned. Actually workload is one of the four areas of discussion that lead a performance presentation into an endless debate. These "rat holes" and their relative sizes in terms of time consumed are shown in Figure 10.26. Presenting this cartoon at the beginning of a presentation helps to avoid these areas.

FIGURE 10.26 Four issues in performance presentations that commonly lead to endless discussion.

Raj Jain, "The Art of Computer Systems Performance Analysis," Wiley, 1991.



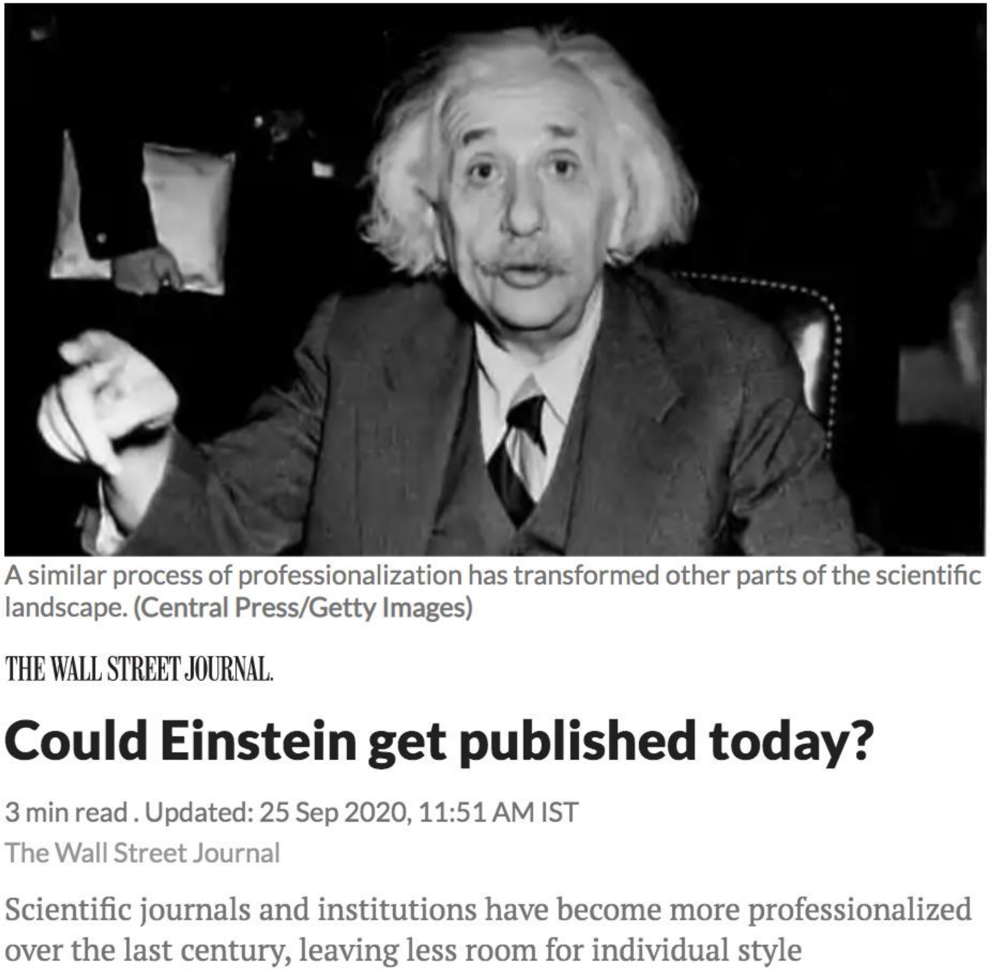
Box 10.2 Reasons for Not Accepting the Results of an Analysis

1. This needs more analysis.
2. You need a better understanding of the workload.
3. It improves performance only for long I/O's, packets, jobs, and files, and most of the I/O's, packets, jobs, and files are short.
4. It improves performance only for short I/O's, packets, jobs, and files, but who cares for the performance of short I/O's, packets, jobs, and files; its the long ones that impact the system.
5. It needs too much memory/CPU/bandwidth and memory/CPU/bandwidth isn't free.
6. It only saves us memory/CPU/bandwidth and memory/CPU/bandwidth is cheap.
7. There is no point in making the networks (similarly, CPUs/disks/...) faster; our CPUs/disks (any component other than the one being discussed) aren't fast enough to use them.
8. It improves the performance by a factor of *x*, but it doesn't really matter at the user level because everything else is so slow.
9. It is going to increase the complexity and cost.
10. Let us keep it simple stupid (and your idea is not stupid).
11. It is not simple. (Simplicity is in the eyes of the beholder.)
12. It requires too much state.
13. Nobody has ever done that before. (You have a new idea.)
14. It is not going to raise the price of our stock by even an eighth. (Nothing ever does, except rumors.)
15. This will violate the IEEE, ANSI, CCITT, or ISO standard.
16. It may violate some future standard.
17. The standard says nothing about this and so it must not be important.
18. Our competitors don't do it. If it was a good idea, they would have done it.
19. Our competition does it this way and you don't make money by copying others.
20. It will introduce randomness into the system and make debugging difficult.
21. It is too deterministic; it may lead the system into a cycle.
22. It's not interoperable.
23. This impacts hardware.
24. That's beyond today's technology.
25. It is not self-stabilizing.
26. Why change—it's working OK.

Raj Jain, "The Art of Computer Systems Performance Analysis," Wiley, 1991.

A Fun Reading: Food for Thought

- https://www.livemint.com/science/news/could-einstein-get-published-today-11601014633853.html



A similar process of professionalization has transformed other parts of the scientific landscape. (Central Press/Getty Images)

THE WALL STREET JOURNAL.

Could Einstein get published today?

3 min read . Updated: 25 Sep 2020, 11:51 AM IST

The Wall Street Journal

Scientific journals and institutions have become more professionalized over the last century, leaving less room for individual style

Computer Architecture

Lecture 5: Memory Security and Reliability (in 2020)

Prof. Onur Mutlu

ETH Zürich

Fall 2020

1 October 2020

Backup Slides

Read Disturb in Flash Memory

Many Errors and Their Mitigation [PIEEE'17]

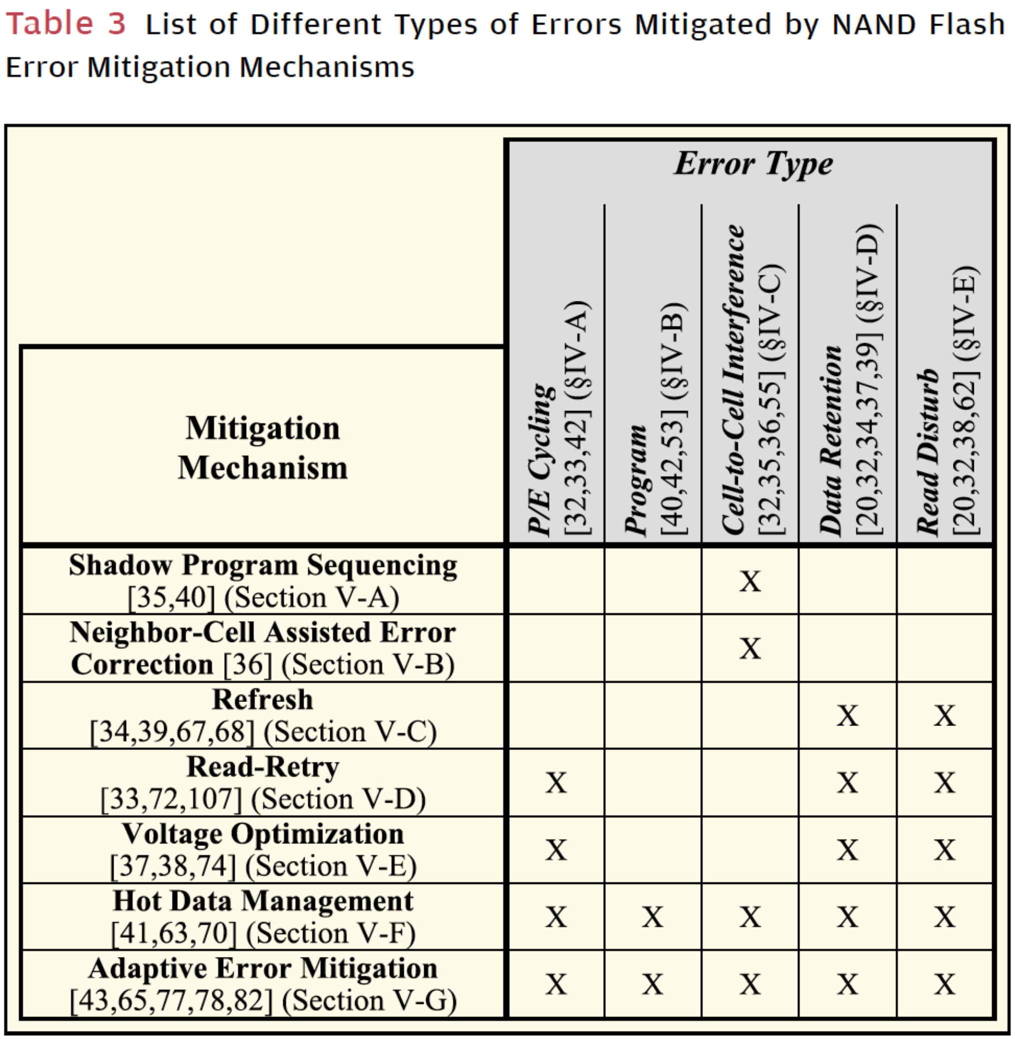


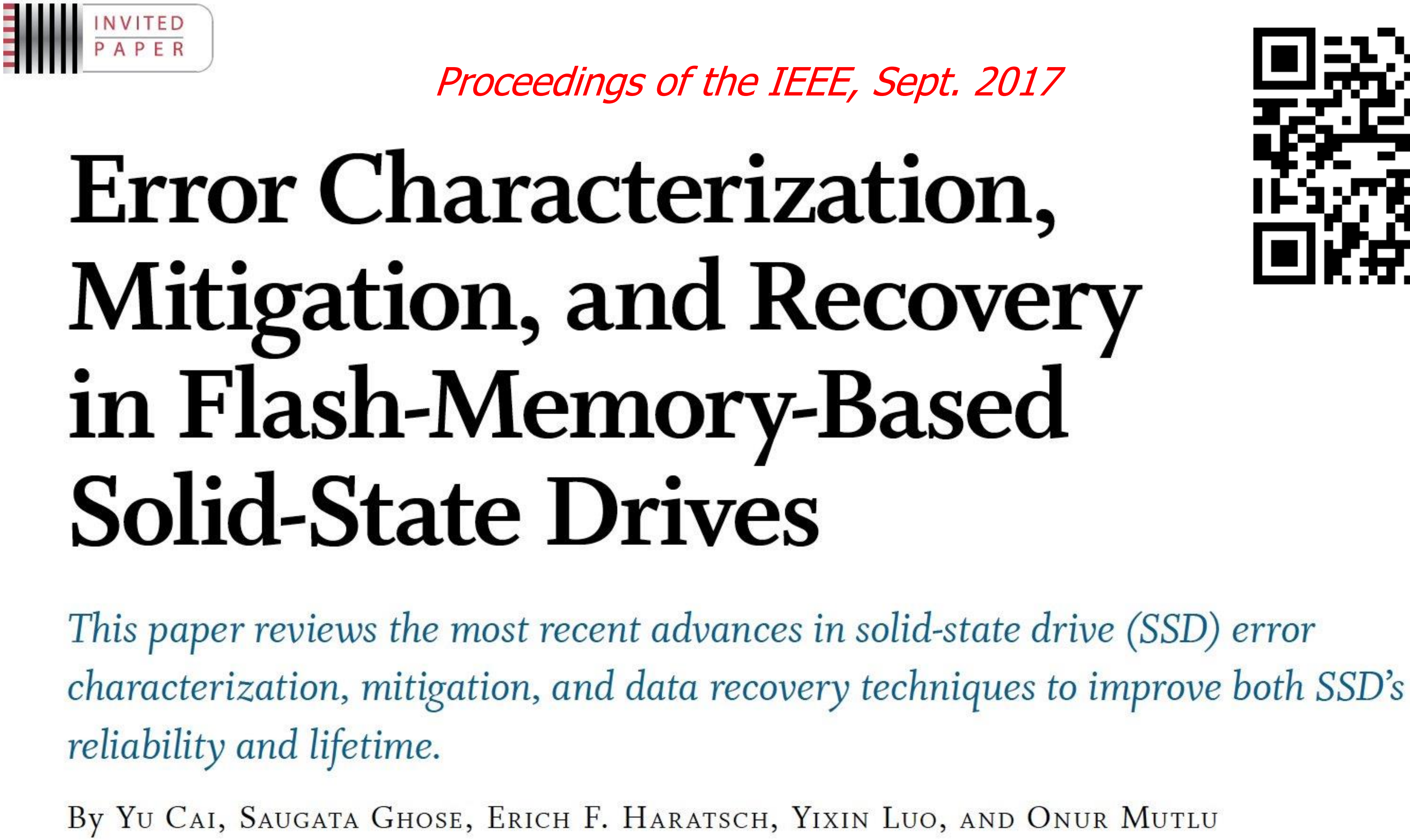
Table 3 List of Different Types of Errors Mitigated by NAND Flash Error Mitigation Mechanisms

	Error Type				
Mitigation Mechanism	***P/E Cycling*** [32,33,42] (§IV-A)	***Program*** [40,42,53] (§IV-B)	***Cell-to-Cell Interference*** [32,35,36,55] (§IV-C)	***Data Retention*** [20,32,34,37,39] (§IV-D)	***Read Disturb*** [20,32,38,62] (§IV-E)
Shadow Program Sequencing [35,40] (Section V-A)			X		
Neighbor-Cell Assisted Error Correction [36] (Section V-B)			X		
Refresh [34,39,67,68] (Section V-C)				X	X
Read-Retry [33,72,107] (Section V-D)	X			X	X
Voltage Optimization [37,38,74] (Section V-E)	X			X	X
Hot Data Management [41,63,70] (Section V-F)	X	X	X	X	X
Adaptive Error Mitigation [43,65,77,78,82] (Section V-G)	X	X	X	X	X

Cai+, "Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives," Proc. IEEE 2017.

Many Errors and Their Mitigation [PIEEE'17]

Proceedings of the IEEE, Sept. 2017



INVITED PAPER

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

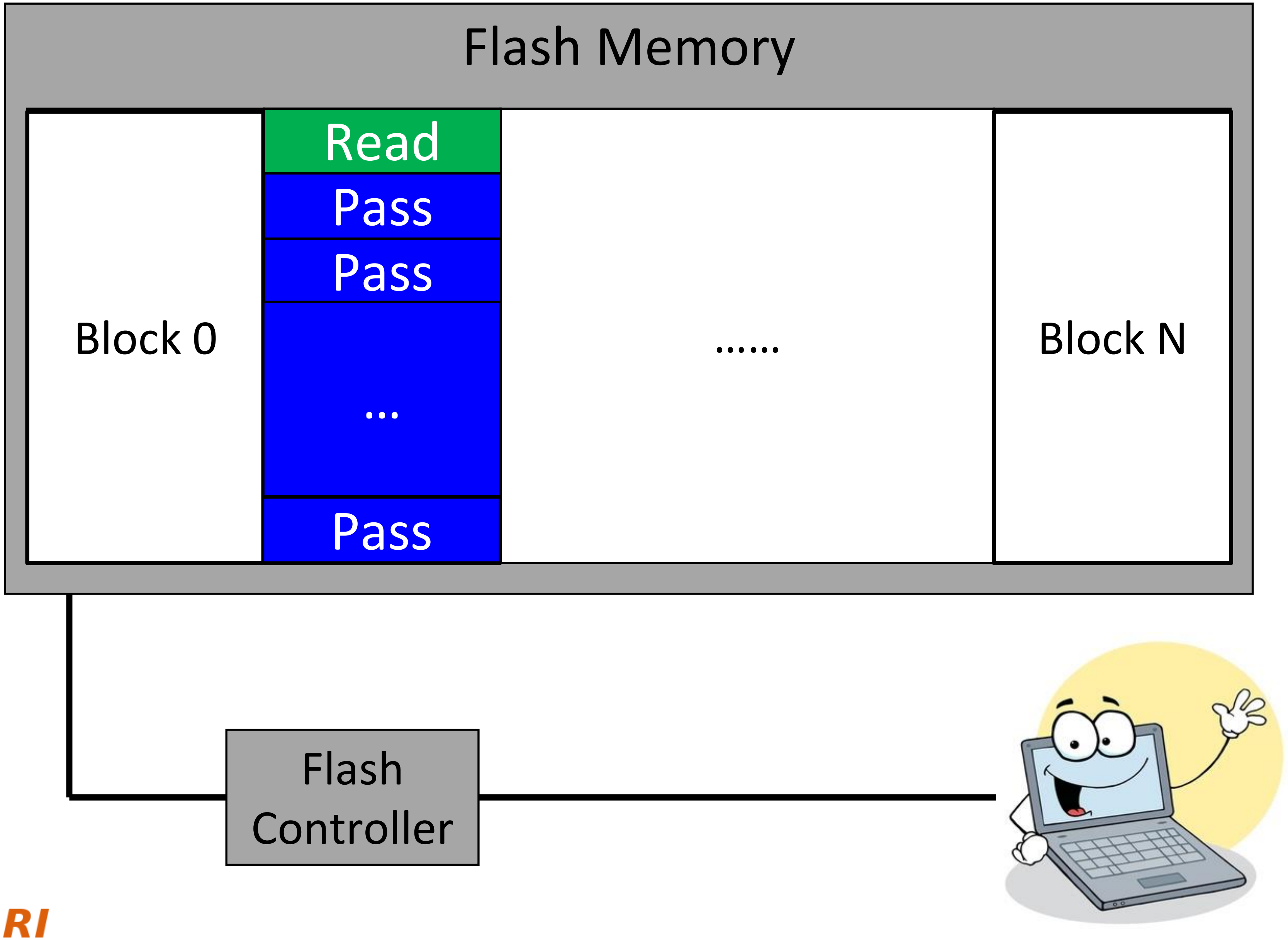
https://arxiv.org/pdf/1706.08642

One Issue: Read Disturb in Flash Memory

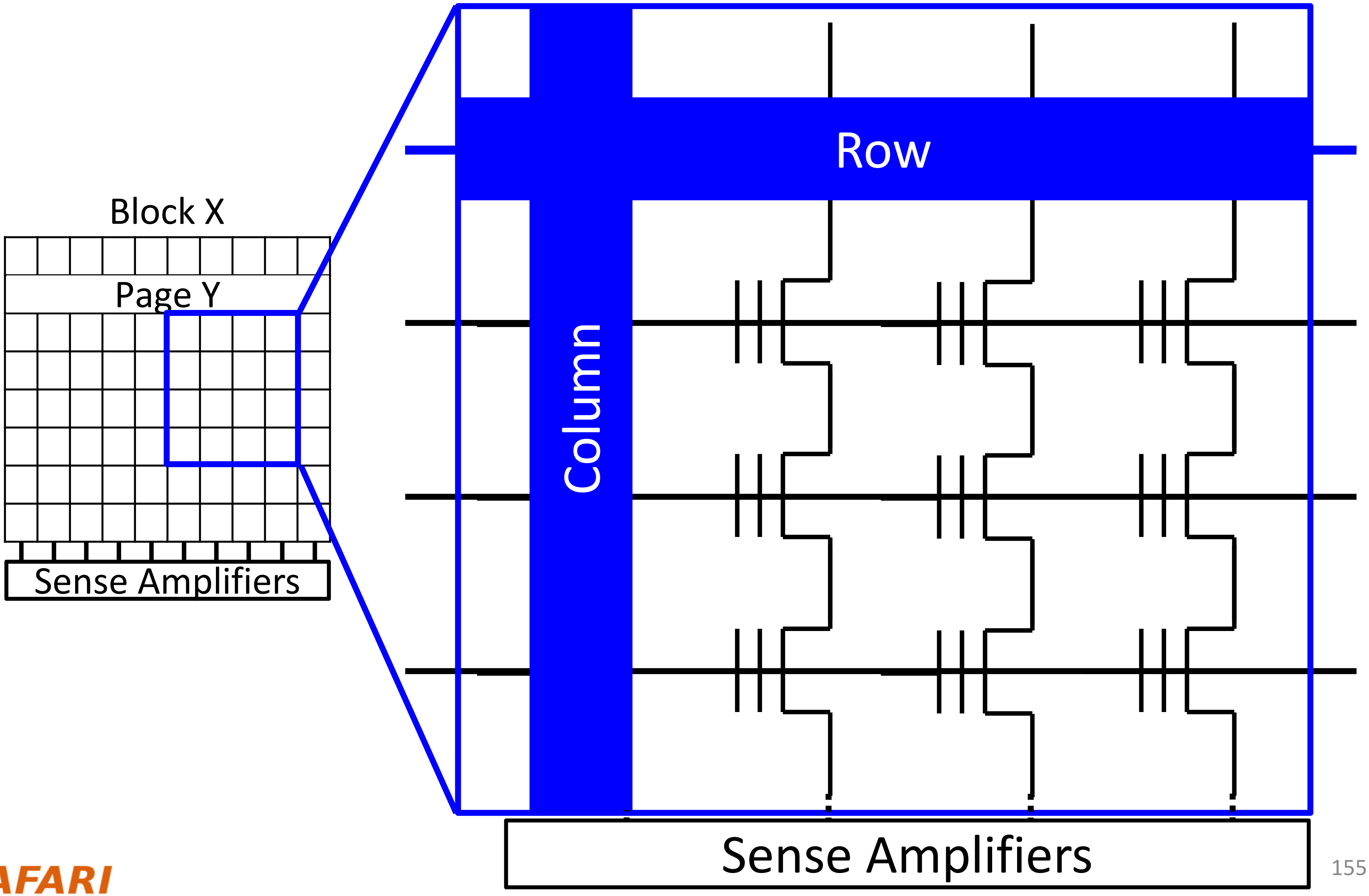
- All scaled memories are prone to read disturb errors

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

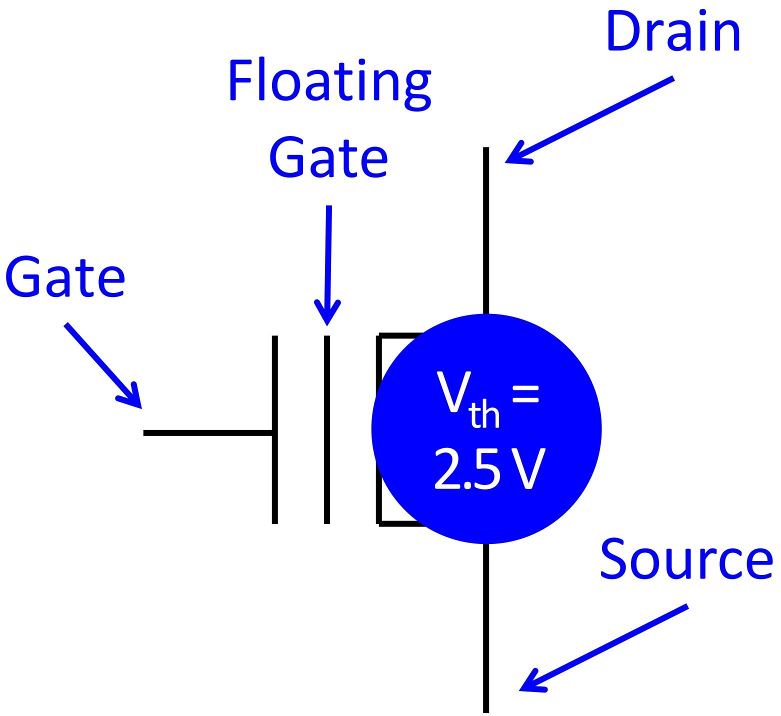
NAND Flash Memory Background



Flash Cell Array

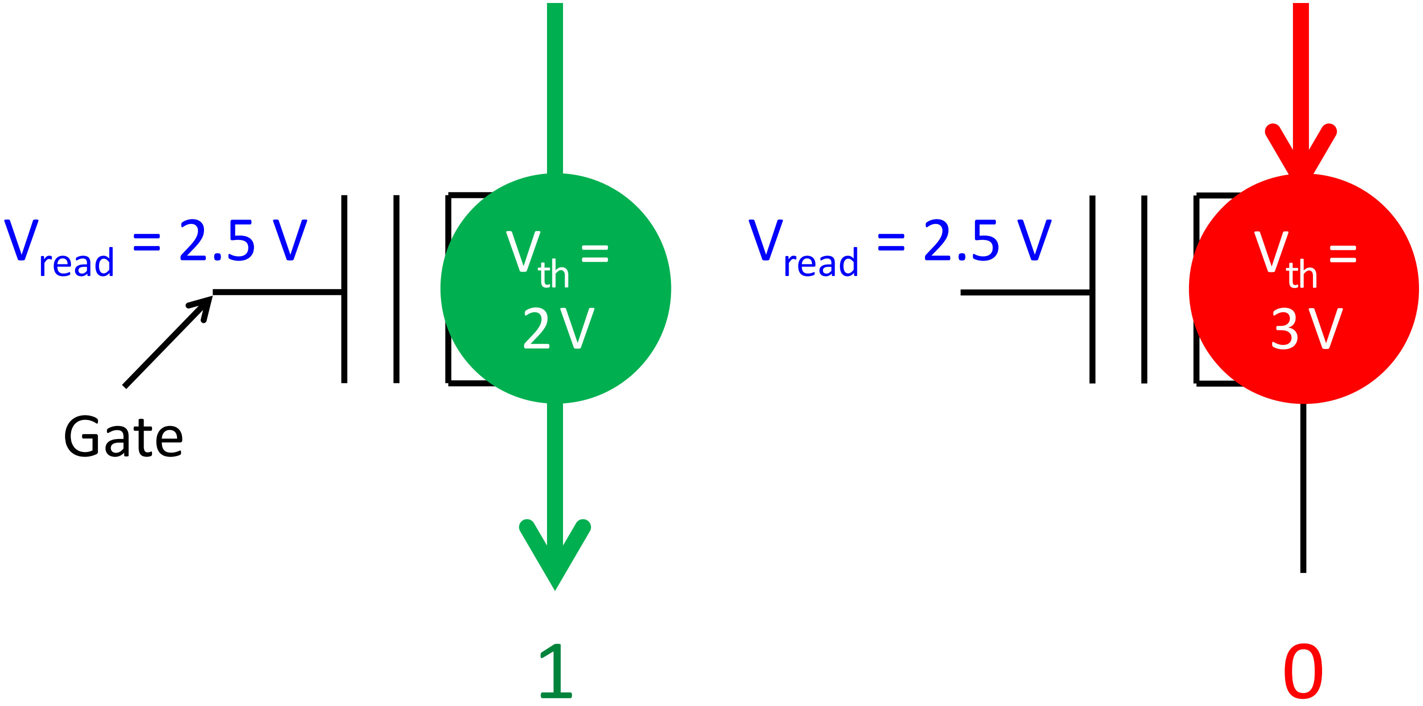


Flash Cell

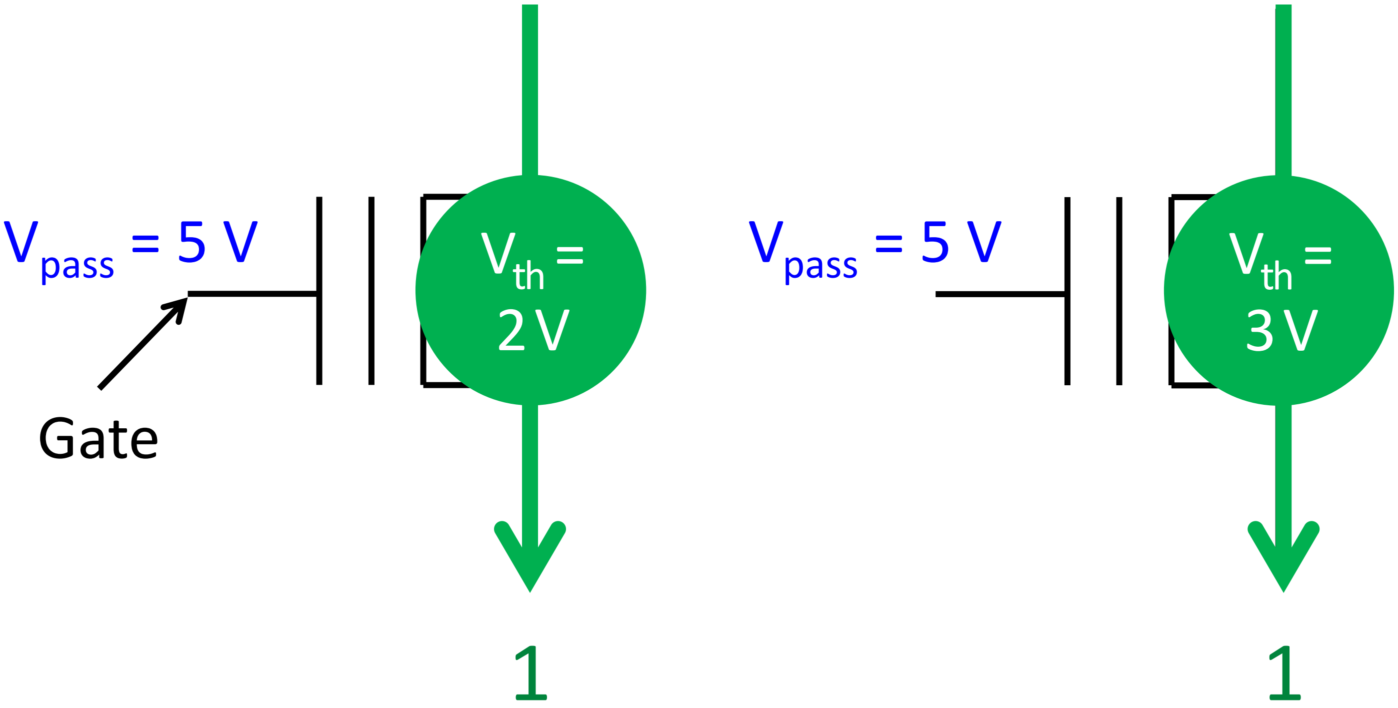


Floating Gate Transistor
(Flash Cell)

Flash Read

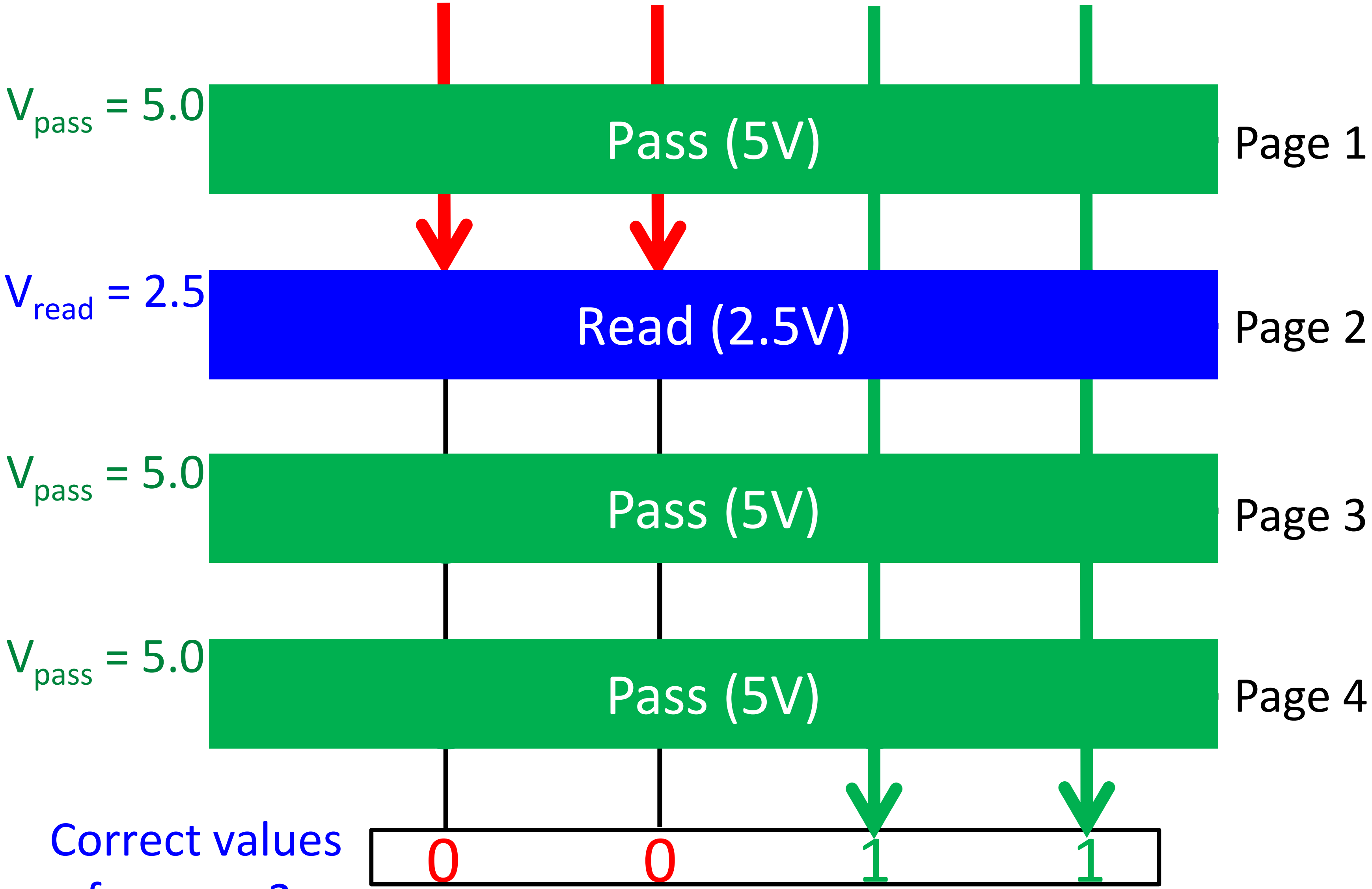


Flash Pass-Through



SAFARI

Read from Flash Cell Array



V_{pass} = 5.0

Pass (5V)

Page 1

V_{read} = 2.5

Read (2.5V)

Page 2

V_{pass} = 5.0

Pass (5V)

Page 3

V_{pass} = 5.0

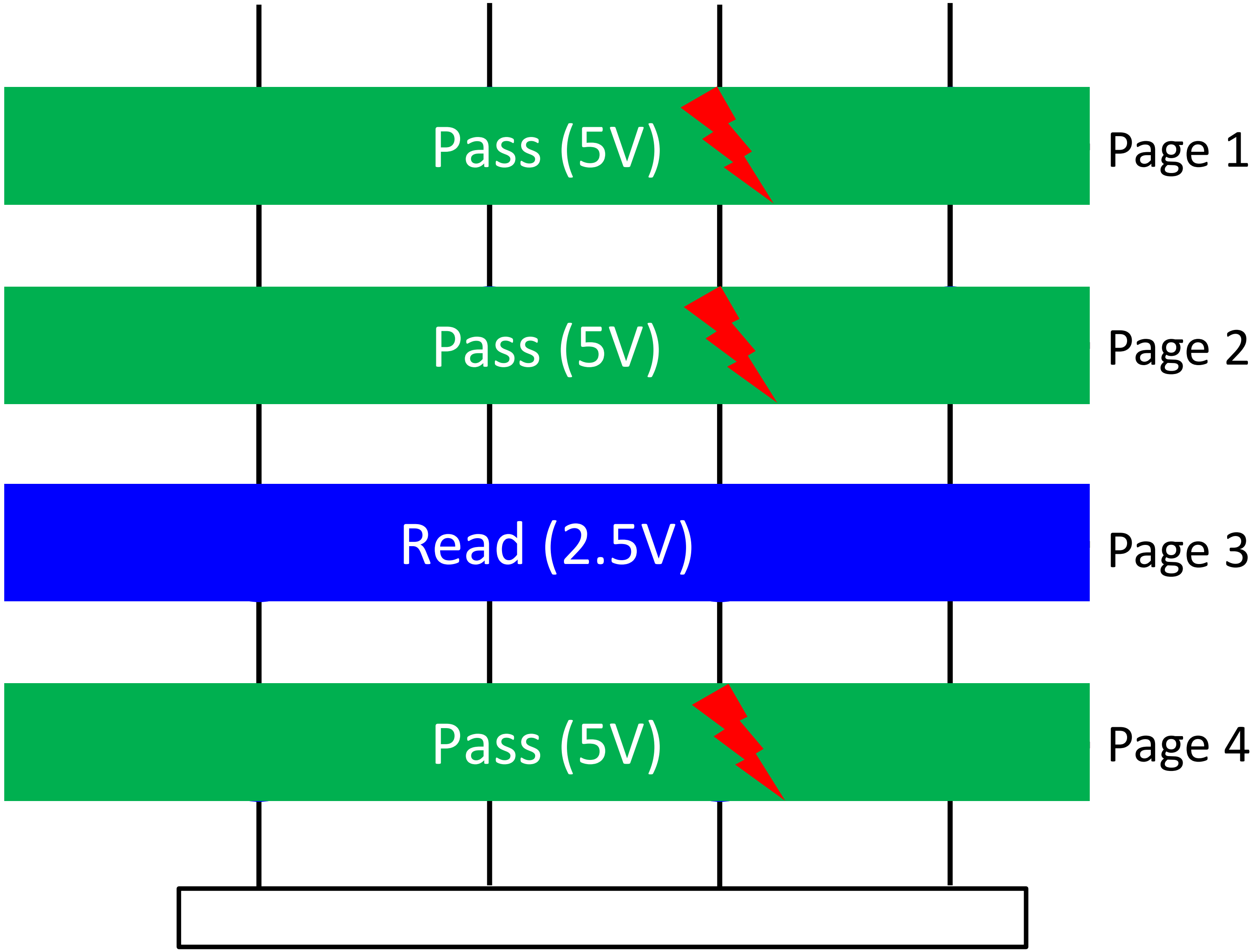
Pass (5V)

Page 4

Correct values for page 2:

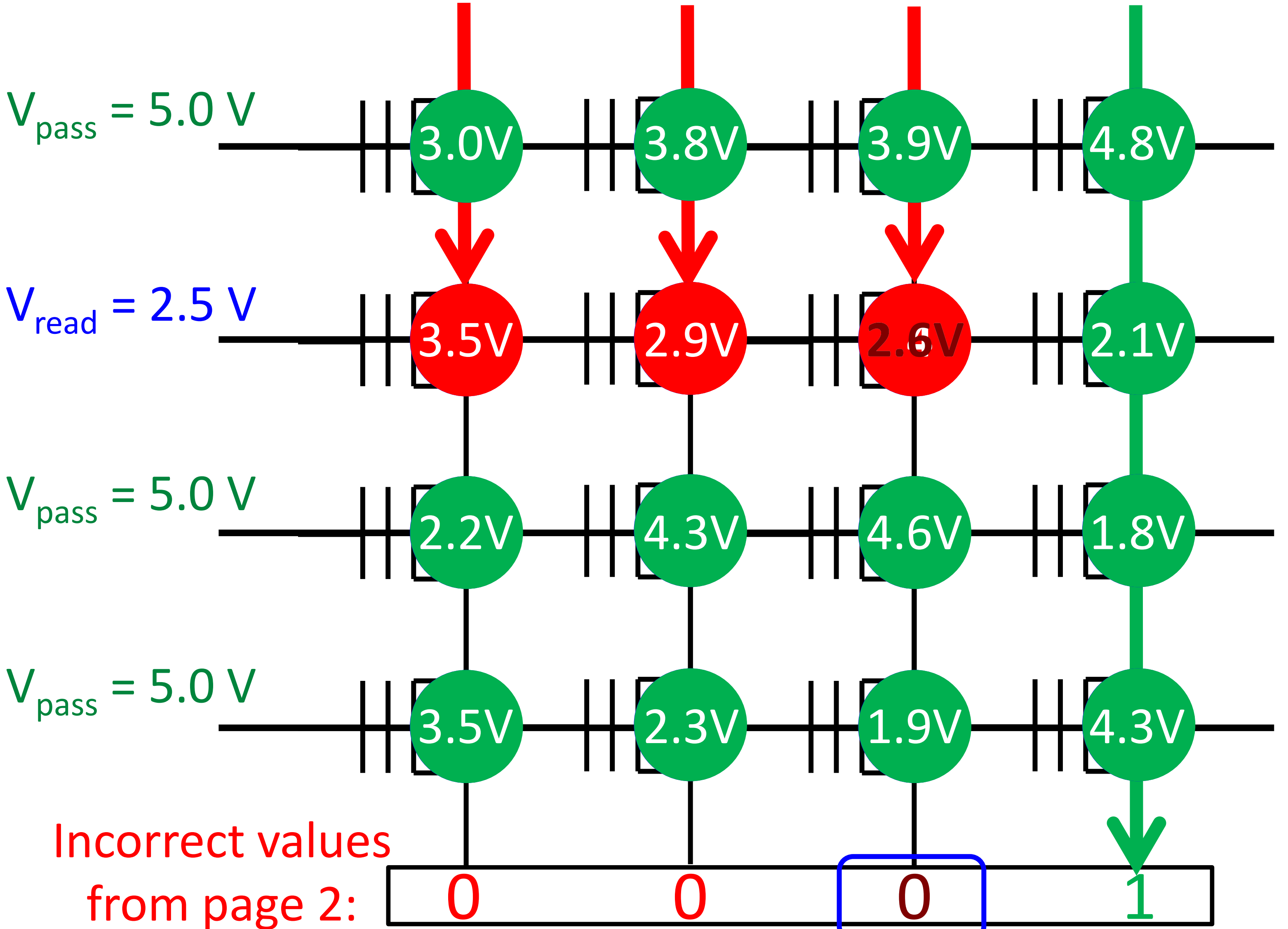
0	0	1	1

Read Disturb Problem: “Weak Programming” Effect



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

Read Disturb Problem: “Weak Programming” Effect



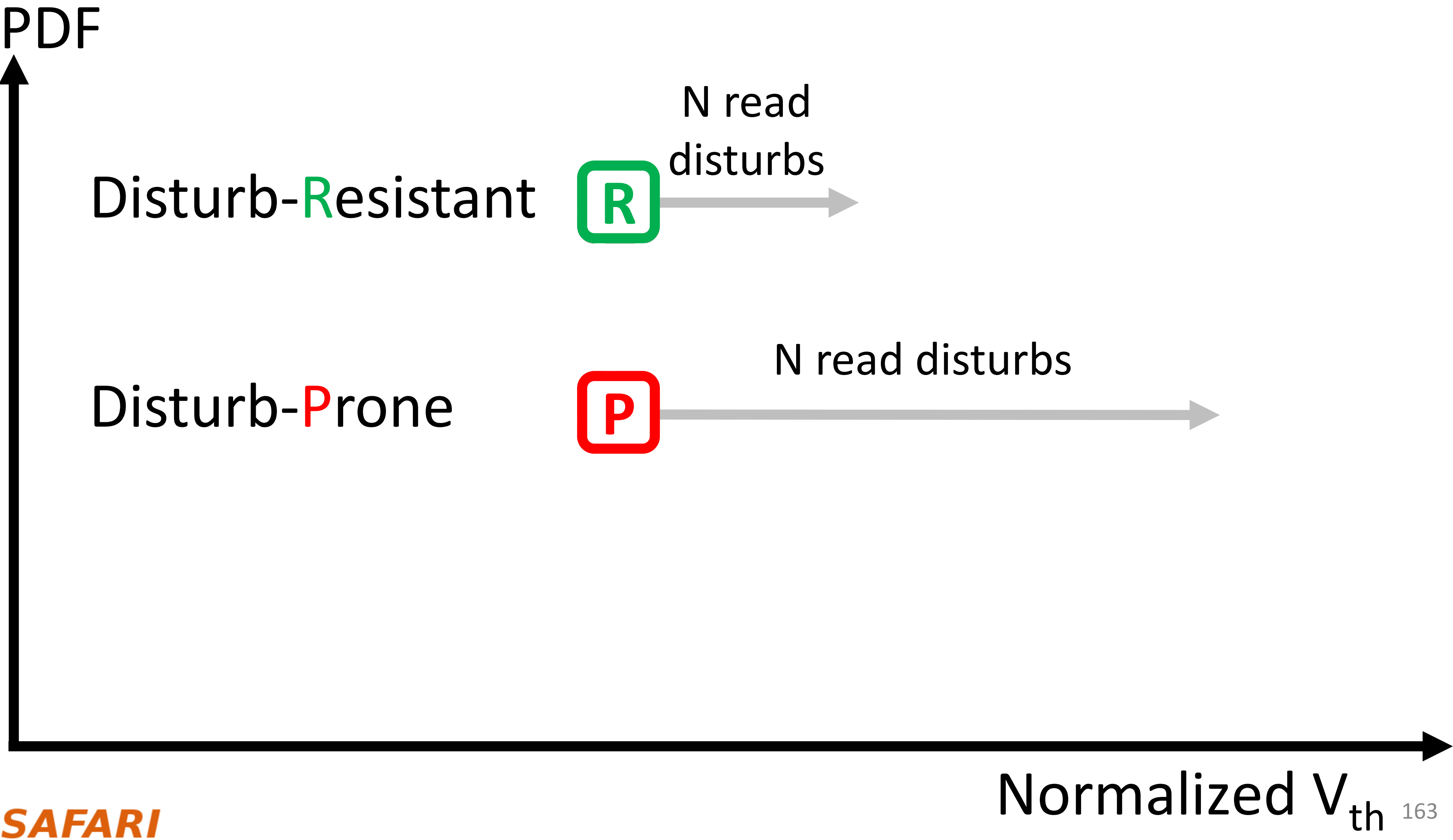
Incorrect values from page 2:

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%

Read Disturb Prone vs. Resistant Cells

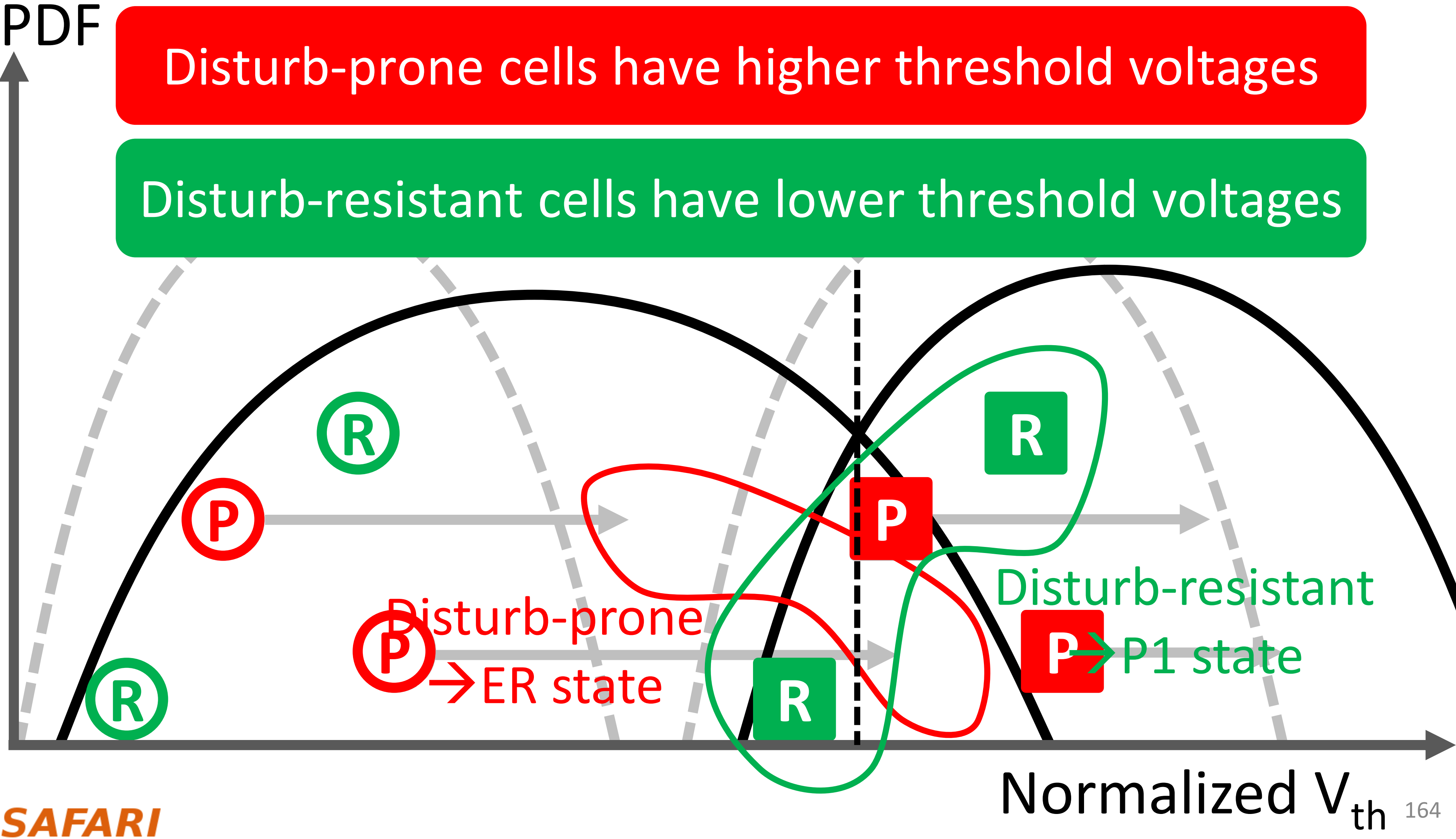


Observation 2: Some Flash Cells Are More Prone to Read Disturb

After 250K read disturbs:

Disturb-prone cells have higher threshold voltages

Disturb-resistant cells have lower threshold voltages



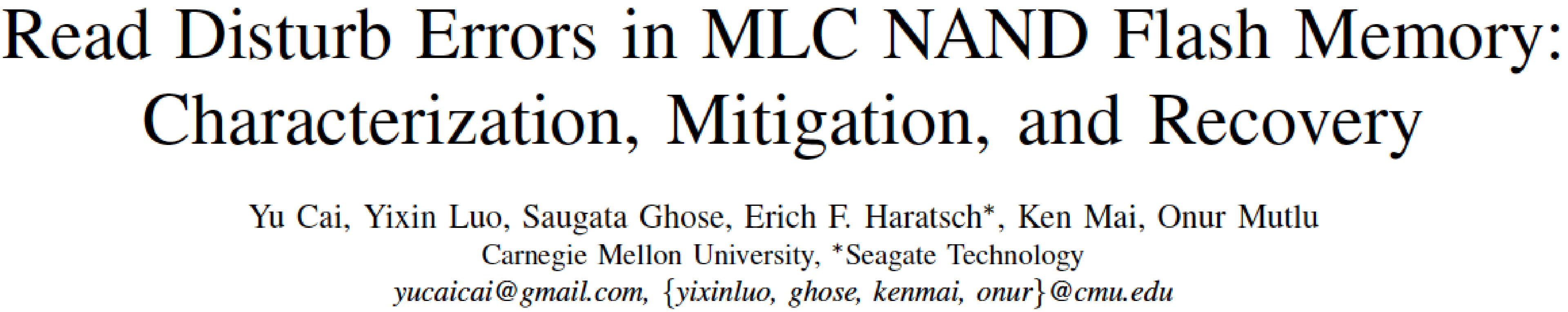
Read Disturb Oriented Error Recovery (RDR)

- 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 → Higher V_{th} state
    - Cells with less V_{th} shifts are disturb-resistant → Lower V_{th} state

Reduces total error count 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.

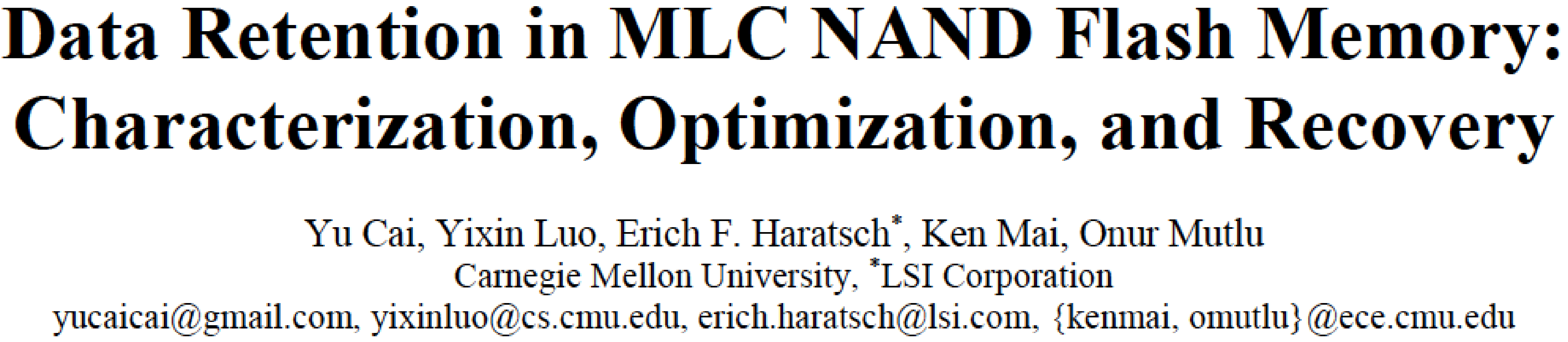


Read Disturb Errors in MLC NAND Flash Memory: Characterization, Mitigation, and Recovery

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

Data Retention in Flash Memory

- 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)]



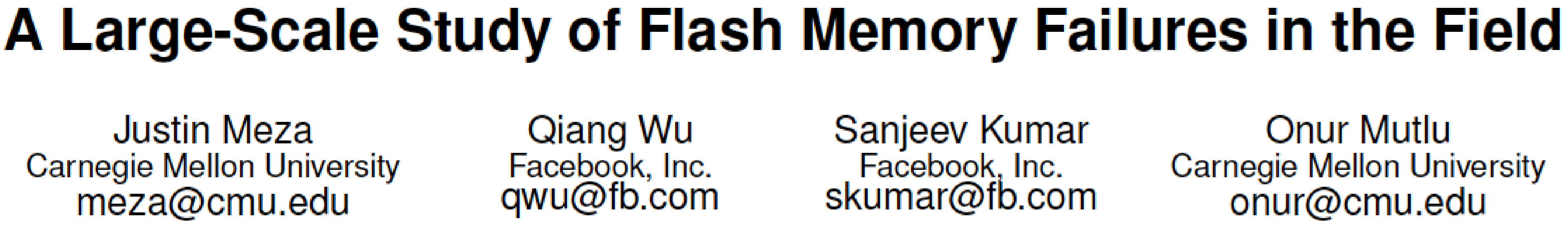
Data Retention in MLC NAND Flash Memory: Characterization, Optimization, and Recovery

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

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

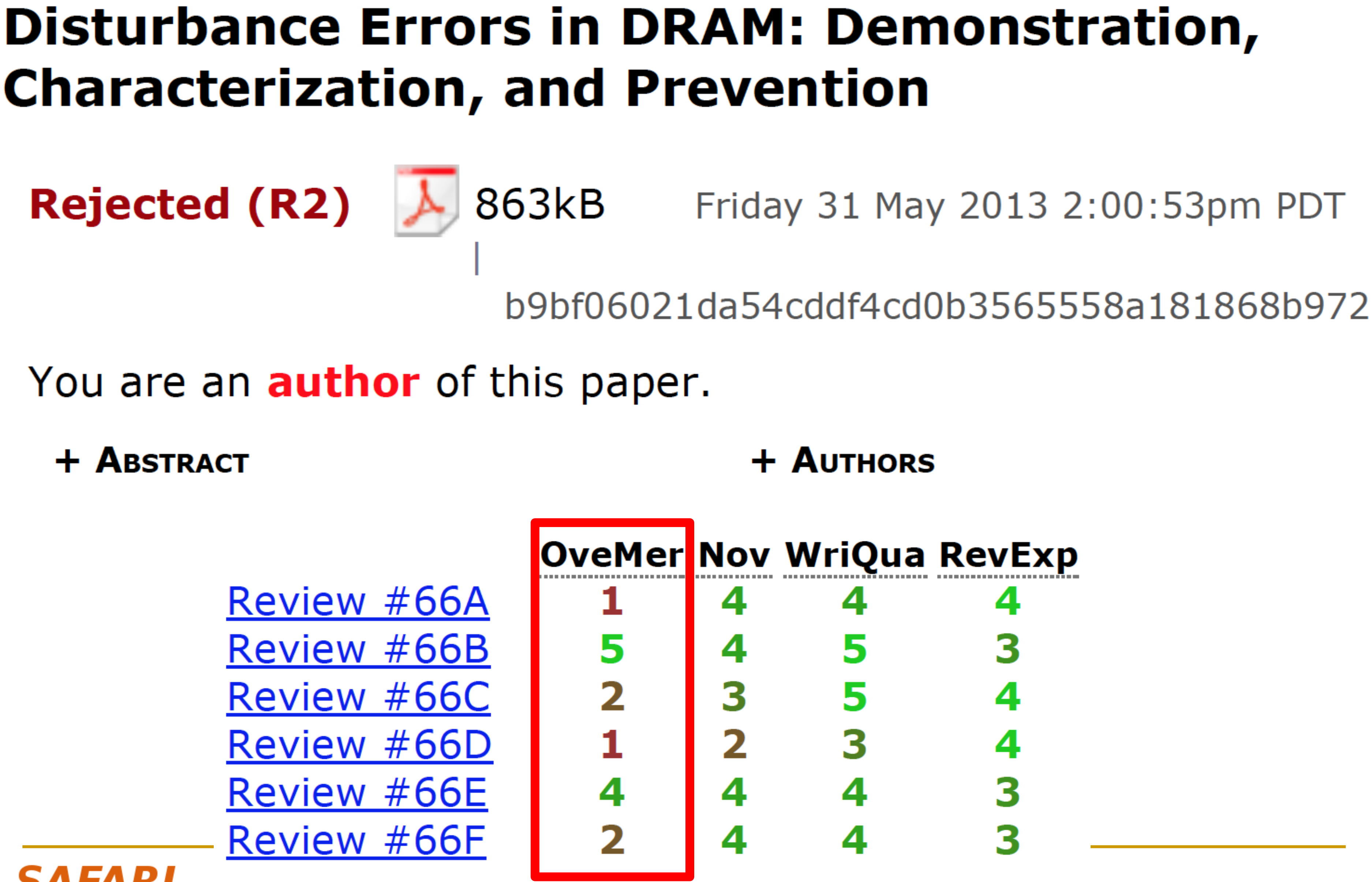
Qiang Wu
Facebook, Inc.
qwu@fb.com

Sanjeev Kumar
Facebook, Inc.
skumar@fb.com

Onur Mutlu
Carnegie Mellon University
onur@cmu.edu

Other Backup Slides

Initial RowHammer Reviews



Disturbance Errors in DRAM: Demonstration, Characterization, and Prevention

Rejected (R2) 863kB | Friday 31 May 2013 2:00:53pm PDT b9bf06021da54cddf4cd0b3565558a181868b972

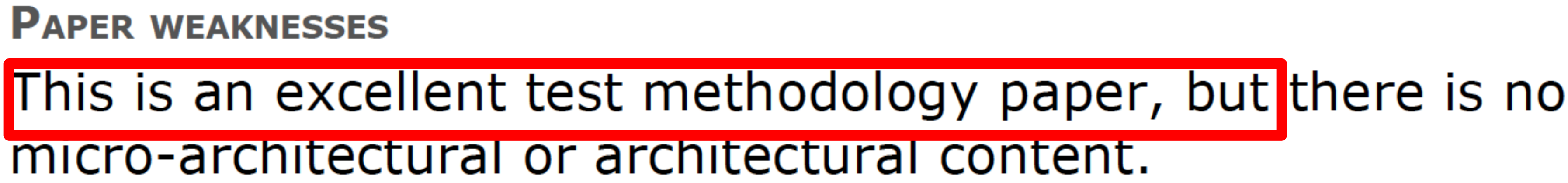
You are an **author** of this paper.

+ Abstract + Authors

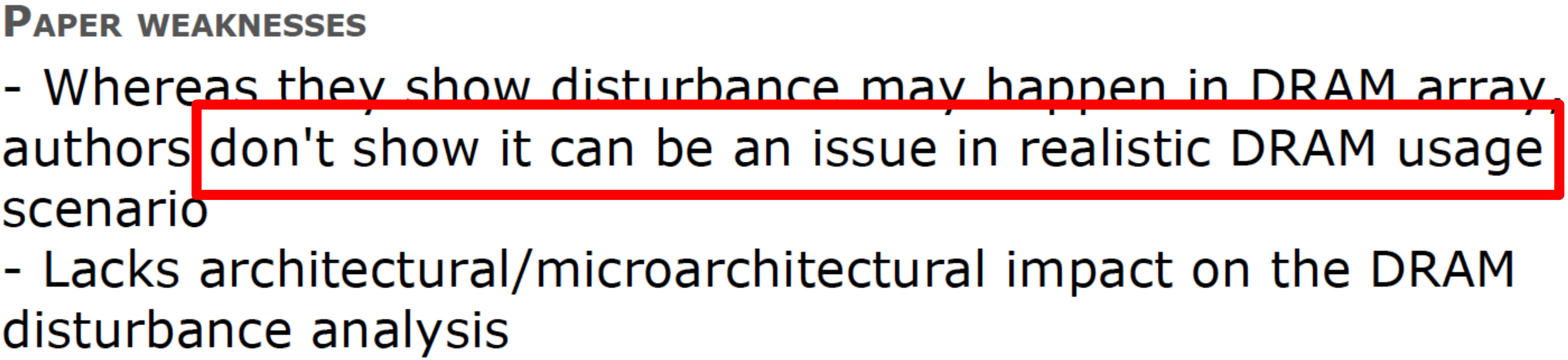
	OveMer	Nov	WriQua	RevExp
Review #66A	1	4	4	4
Review #66B	5	4	5	3
Review #66C	2	3	5	4
Review #66D	1	2	3	4
Review #66E	4	4	4	3
Review #66F	2	4	4	3

Missing the Point

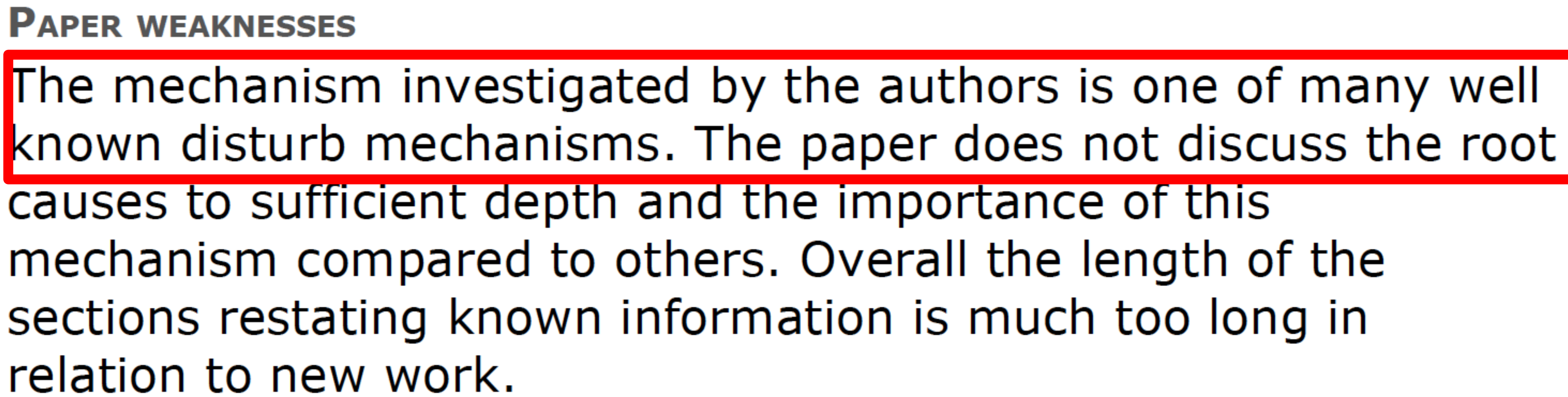
Reviews from Micro 2013


PAPER WEAKNESSES

This is an excellent test methodology paper, but there is no micro-architectural or architectural content.


PAPER WEAKNESSES

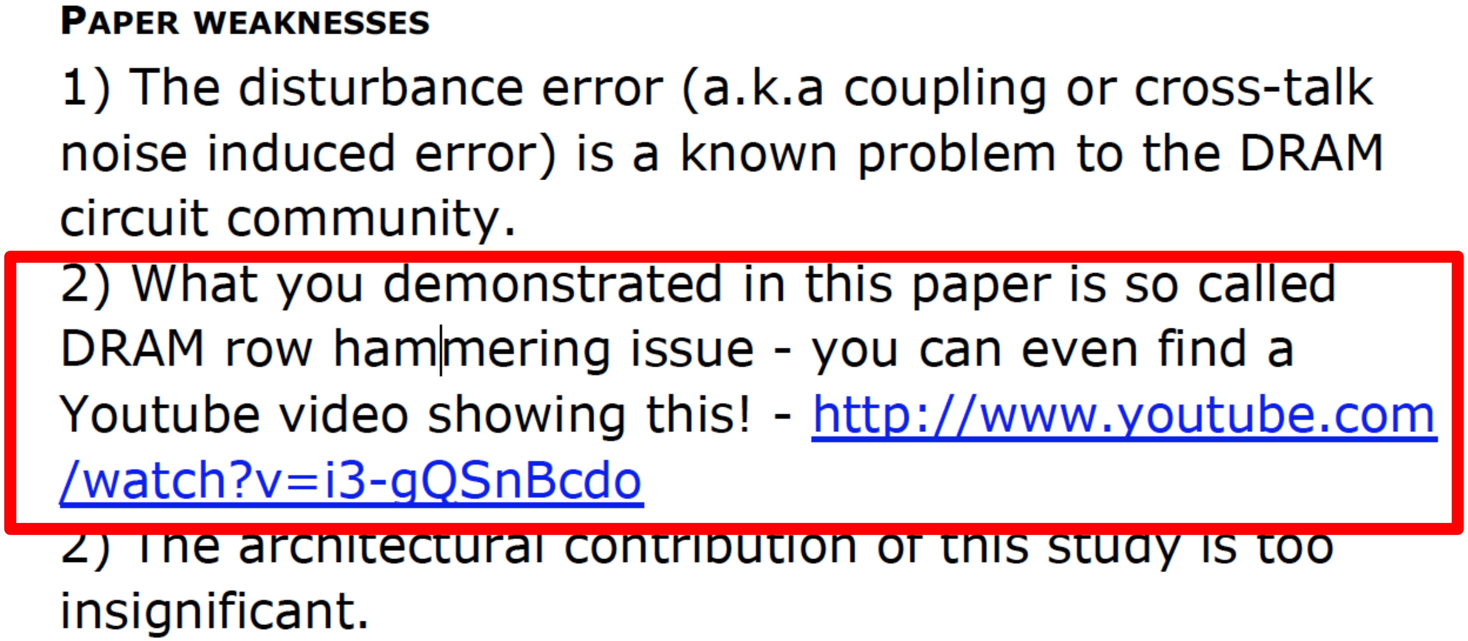
- Whereas they show disturbance may happen in DRAM array, authors don't show it can be an issue in realistic DRAM usage scenario
- Lacks architectural/microarchitectural impact on the DRAM disturbance analysis


PAPER WEAKNESSES

The mechanism investigated by the authors is one of many well known disturb mechanisms. The paper does not discuss the root causes to sufficient depth and the importance of this mechanism compared to others. Overall the length of the sections restating known information is much too long in relation to new work.

More …

Reviews from ISCA 2014

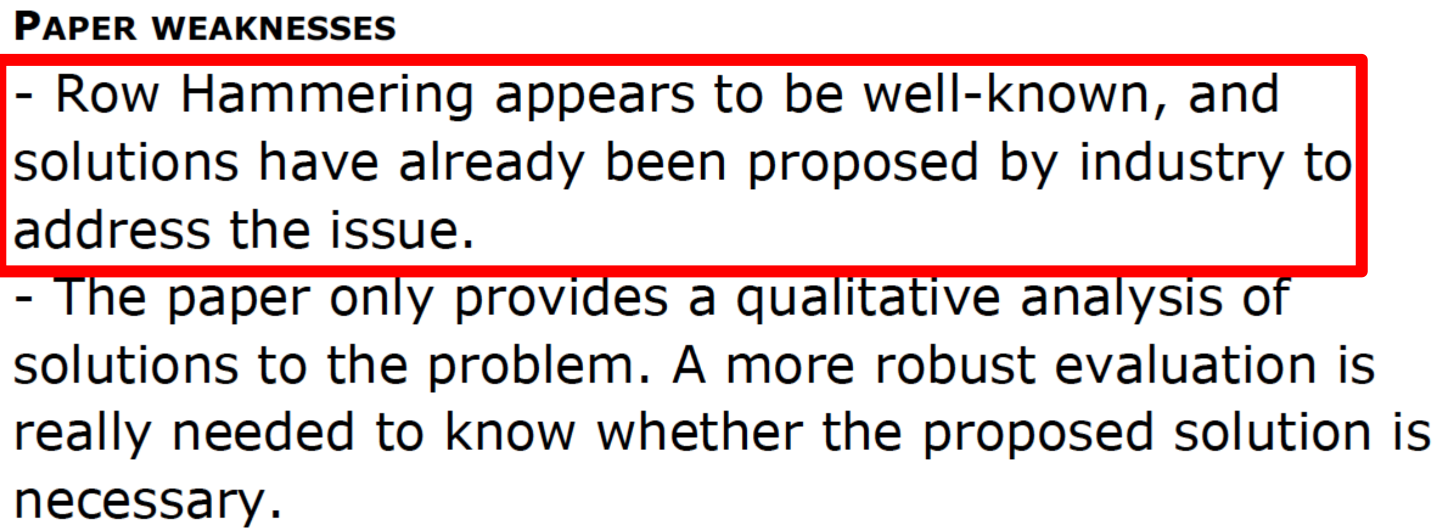


PAPER WEAKNESSES

1) The disturbance error (a.k.a coupling or cross-talk noise induced error) is a known problem to the DRAM circuit community.

2) What you demonstrated in this paper is so called DRAM row hammering issue - you can even find a Youtube video showing this! - http://www.youtube.com/watch?v=i3-gQSnBcdo

2) The architectural contribution of this study is too insignificant.

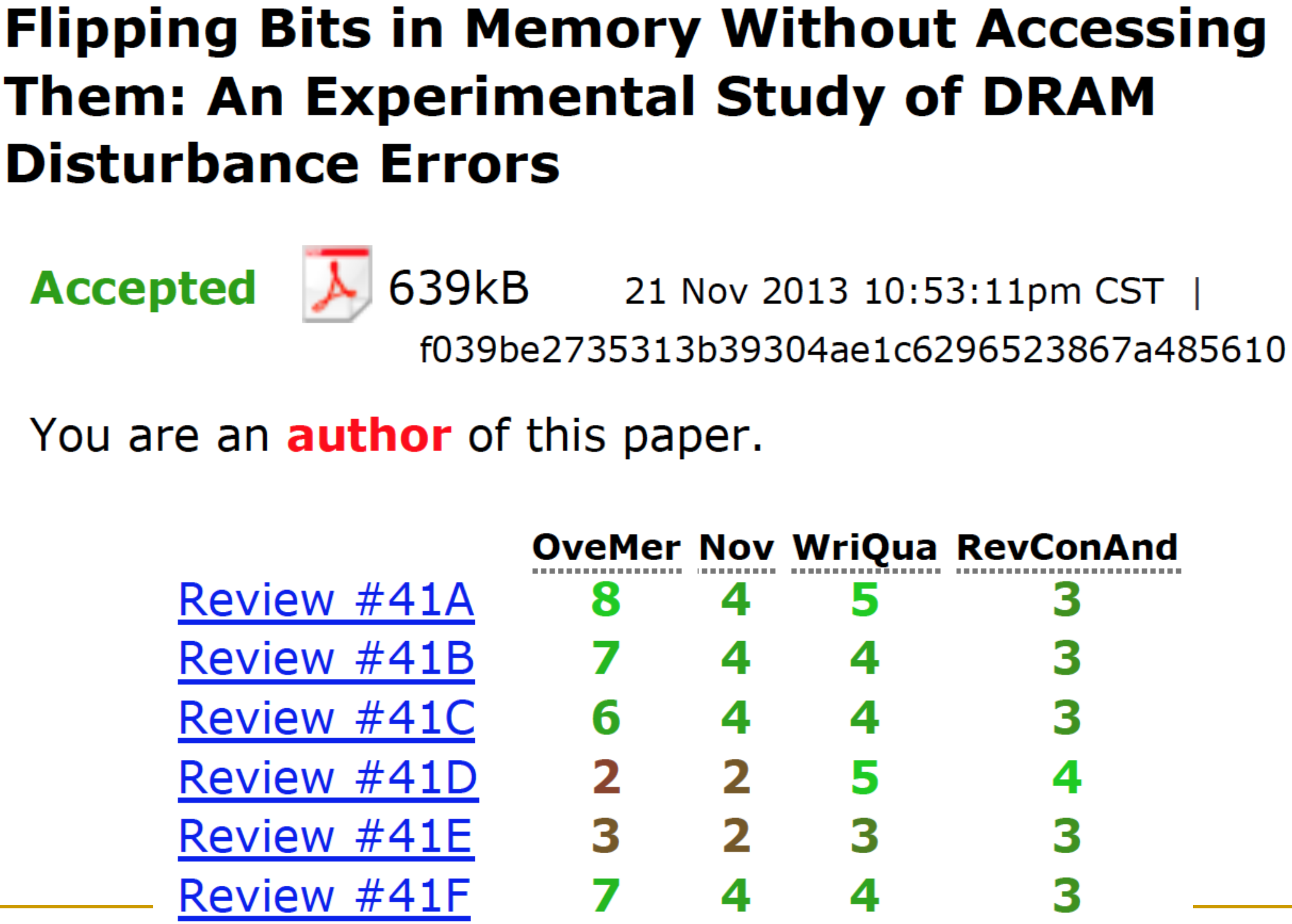


PAPER WEAKNESSES

- Row Hammering appears to be well-known, and solutions have already been proposed by industry to address the issue.

- The paper only provides a qualitative analysis of solutions to the problem. A more robust evaluation is really needed to know whether the proposed solution is necessary.

Final RowHammer Reviews



Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors

Accepted 639kB 21 Nov 2013 10:53:11pm CST | f039be2735313b39304ae1c6296523867a485610

You are an **author** of this paper.

	OveMer	Nov	WriQua	RevConAnd
Review #41A	8	4	5	3
Review #41B	7	4	4	3
Review #41C	6	4	4	3
Review #41D	2	2	5	4
Review #41E	3	2	3	3
Review #41F	7	4	4	3