Enabling Efficient and Scalable Read Disturbance Mitigation

via New Experimental Insights into Modern Memory Chips

agyaglikci.github.io

Abdullah Giray Yaglikci

agyaglikci@gmail.com

https://agyaglikci.github.io

13 March 2024

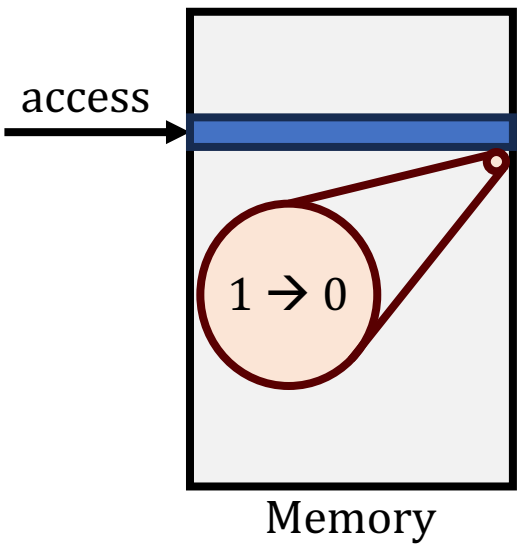
ARM Cambridge

safari.ethz.ch

SAFARI

ETH zürich

Lack of Memory Isolation



Data **Loss** or **Corruption**

Compromise Application **Correctness**

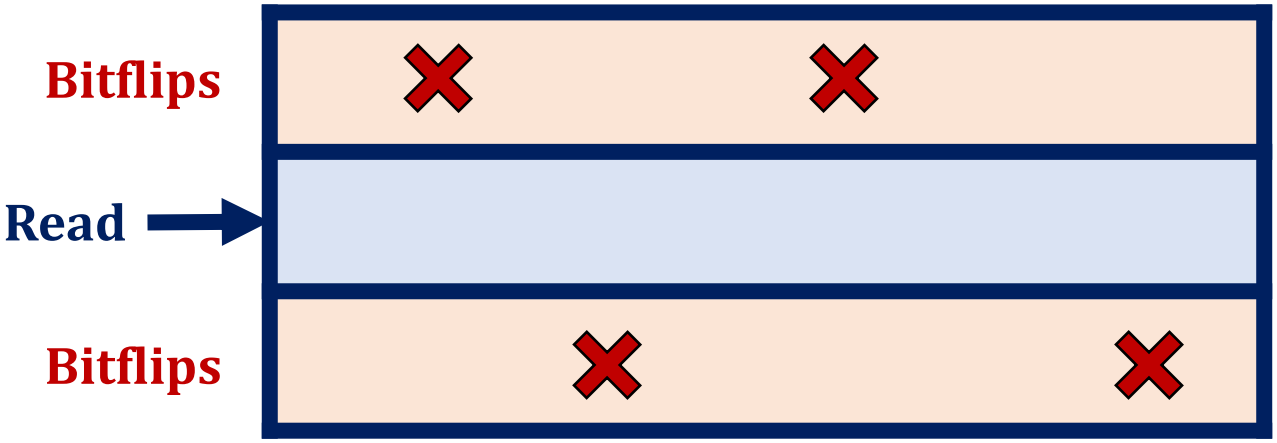
Leak Private Information

Take Over a Computer

An access to one memory address
should not have **unintended side effects**
on data stored in **other addresses**

Memory isolation is **difficult in modern memory chips**

DRAM Read Disturbance

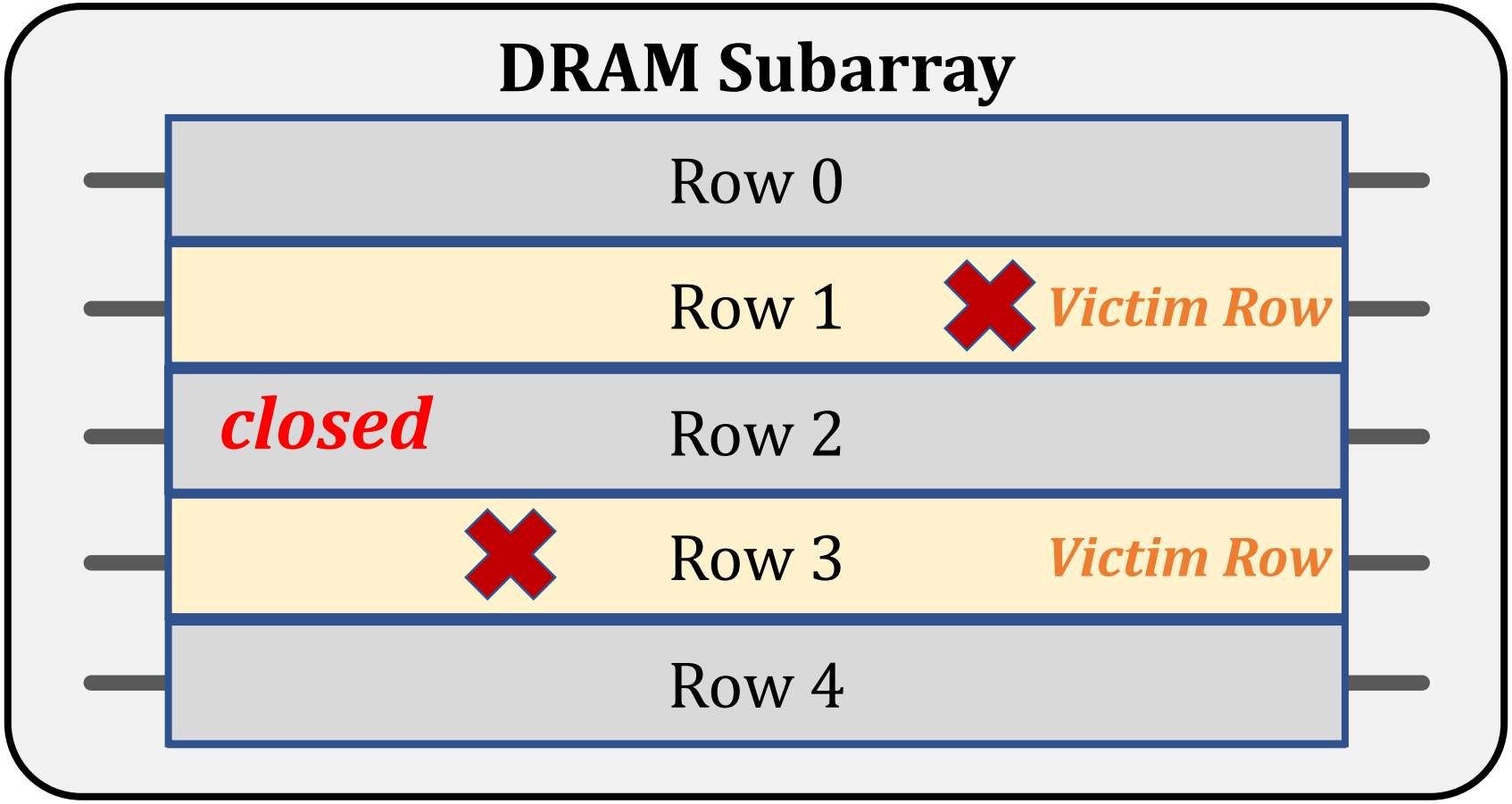


Reading from a memory location
disturbs data in **physically nearby** locations

DRAM Read Disturbance – Swimming Pool Analogy

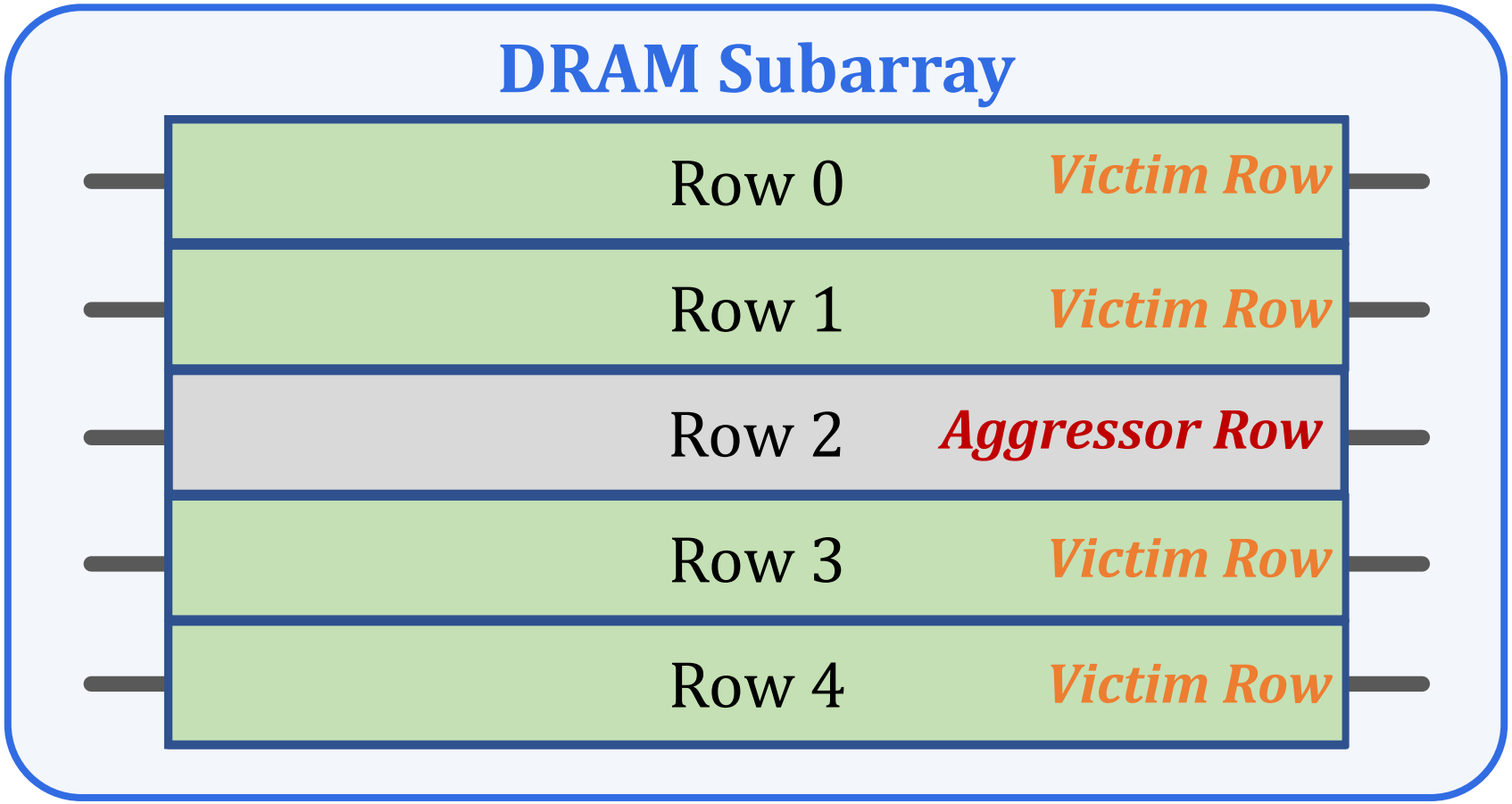
Swimming **in a lane** disturbs **nearby lanes**

The RowHammer Vulnerability [Kim+, ISCA'14]



Repeatedly **opening** (activating) and **closing** (precharging) a DRAM row causes **RowHammer bitflips** in nearby cells and breaks **memory isolation**

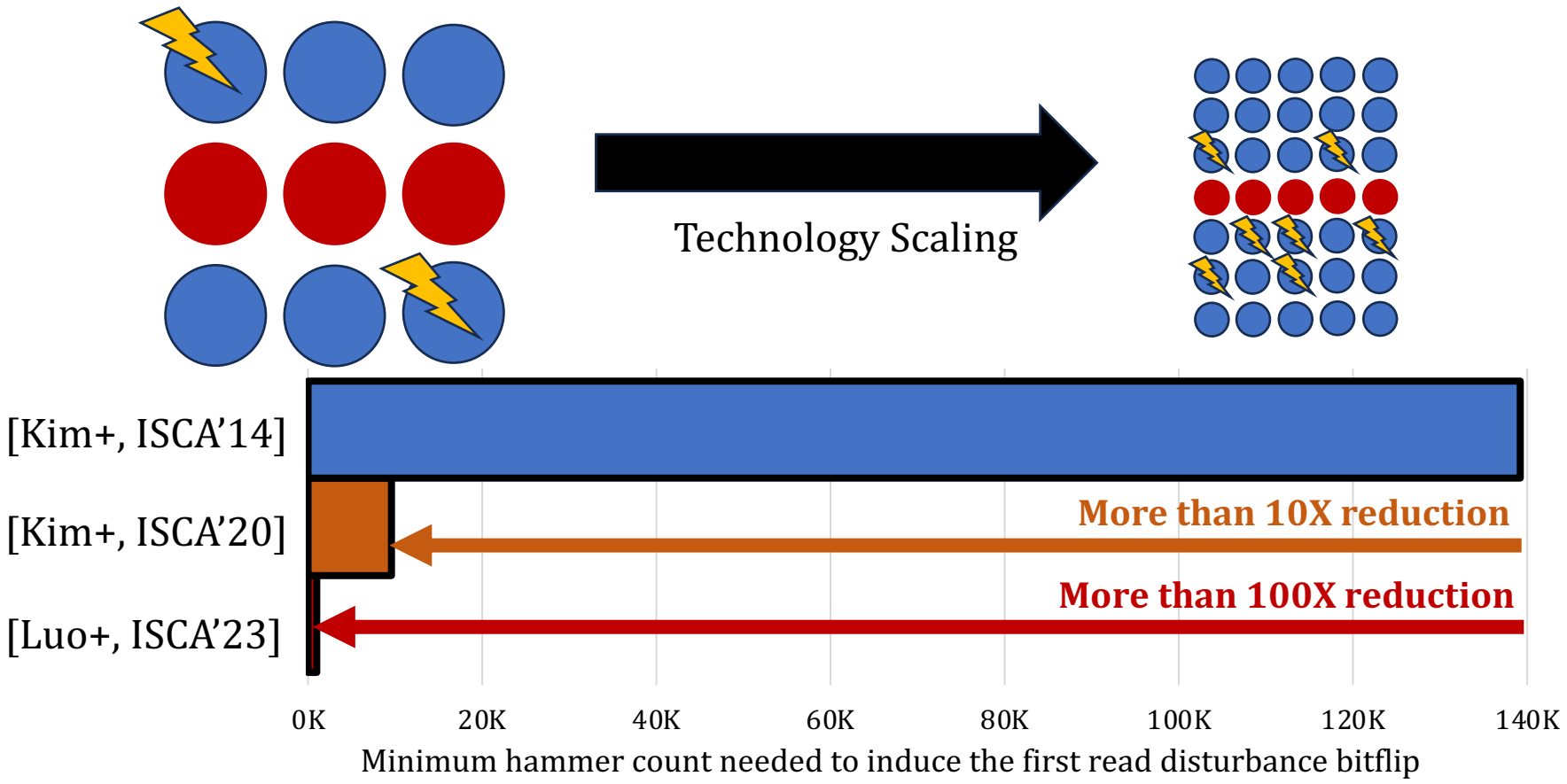
An Example RowHammer Mitigation: Preventive Refresh



Refreshing potential victim rows
mitigates read disturbance bitflips

[Kim+ ISCA'20]

Motivation



DRAM chips are **increasingly more vulnerable** to **read disturbance** with **technology scaling**

Two Main Types of DRAM Refresh

1 **Periodic Refresh:**

Periodically **restores** the charge which DRAM cells leak **over time** **once** in every **32ms** or **64ms**

2 **Preventive Refresh:**

Mitigates read disturbance by **refreshing** **potential victim rows**

Putting into Perspective:

- Bitflips occur at **~1000 row activations** scattered across **64ms**
- 64ms is **as large as** to accommodate **1,280,000 activations**
- An attack needs **only 0.08%** of the **activation budget**
- **1280 rows** can be **concurrently hammered**
- **1280 additional refreshes** are needed

Two Main Types of DRAM Refresh

Periodic Refresh:

An attacker can keep **low profile**
(e.g., uses 0.08% of activation budget)
and **induce bitflips**

Mitigates read disturbance by refreshing potential victim rows

Preventing bitflips requires
tracking many rows and **performing many refreshes**

- **1280 rows** can be **concurrently hammered**
- **1280 additional refreshes** are needed

How Large is 1000 Activations?

- Bitflips occur **at ~1000 activations**
- Mitigation mechanisms trigger **preventive actions** (e.g., preventive refresh) **at ~500 activations**
- Is 500 **a distinctive activation count**?
- Benign workloads activate **hundreds of rows** more than **512 times** in a refresh window

Memory intensive workloads

from SPEC'06/17, TPC, YCSB, and MediaBench

Benchmark	MPKI	# of Rows w/ ACT count >512
429.mcf	68.27	2564
470.lbm	28.09	664
519.lbm	24.37	2482
434.zeusmp	22.24	292
510.parest	17.79	94
437.leslie3d	15.82	7
483.xalancbmk	13.67	113
482.sphinx3	12.59	304
505.mcf	11.35	732
471.omnetpp	10.72	122
tpch2	9.09	88
520.omnetpp	9.00	32
tpch17	7.43	26

Sorted

How Large is 1000 Activations?

- Bitflips occur **at ~1000 activations**
- Mitigation mechanisms trigger **preventive actions** (e.g., preventive refresh) **at ~500 activations**

Memory intensive workloads
from SPEC'06/17, TPC, YCSB, and MediaBench

Benchmark	MPKI	# of Rows w/ ACT count >512
429.mcf	68.27	2564
470.lbm	28.09	664
519.lbm	24.37	2482
434.zeusmp	22.24	292
tpch2	9.09	88
520.omnetpp	9.00	32
tpch17	7.43	26

Benign workloads **might not be so benign**
even if they are **not very memory intensive**

more than **500 times** in a refresh window

Read Disturbance is an Outstanding Problem

Increasing DRAM chip density
exacerbates DRAM read disturbance

Attackers can keep **low profile**
(using <0.16% of the row activation budget)

Benign applications become **not-so-benign**

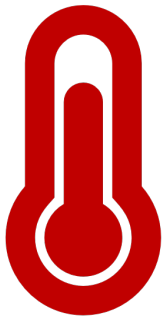
Efficient and scalable solutions are needed

Thesis

A **deeper understanding** of
DRAM read disturbance
is the key to enable
efficient and scalable solutions

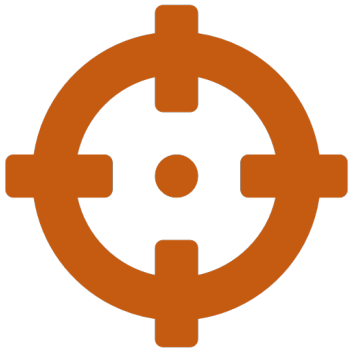
My Dissertation Works

- A deeper look into DRAM read disturbance



Temperature

Memory access patterns



Victim cell's physical location

Voltage

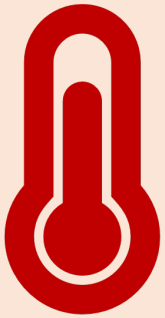
- Solutions to DRAM read disturbance

Leveraging Heterogeneity

Throttling Unsafe Accesses

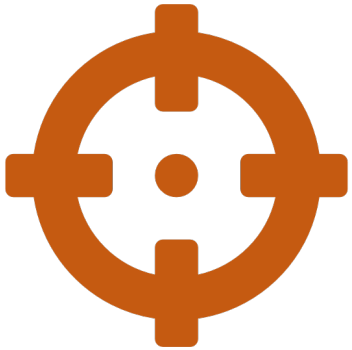
Parallelizing Preventive Actions

Our Recent Works



Temperature

Memory access patterns

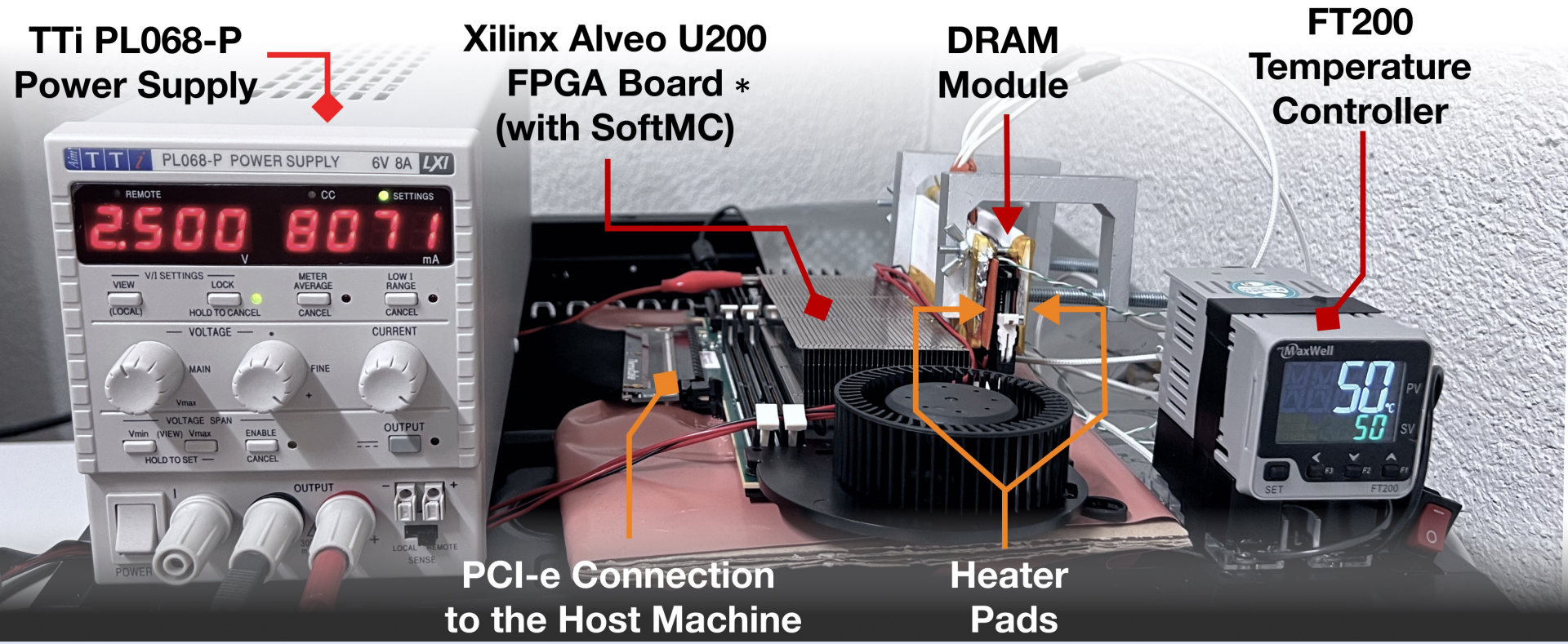


Victim cell's physical location

Leveraging Heterogeneity

DRAM Testing Infrastructure

DRAM Bender on a Xilinx Virtex UltraScale+ XCU200



Fine-grained control over **DRAM commands**, **timing parameters (±1.5ns)**, **temperature (±0.5°C)**, and **wordline voltage (±1mV)**

*Olgun et al., "DRAM Bender: An Extensible and Versatile FPGA-based Infrastructure to Easily Test State-of-the-art DRAM Chips," in TCAD, 2023. [GitHub: https://github.com/CMU-SAFARI/DRAM-Bender]

DRAM Chips Tested

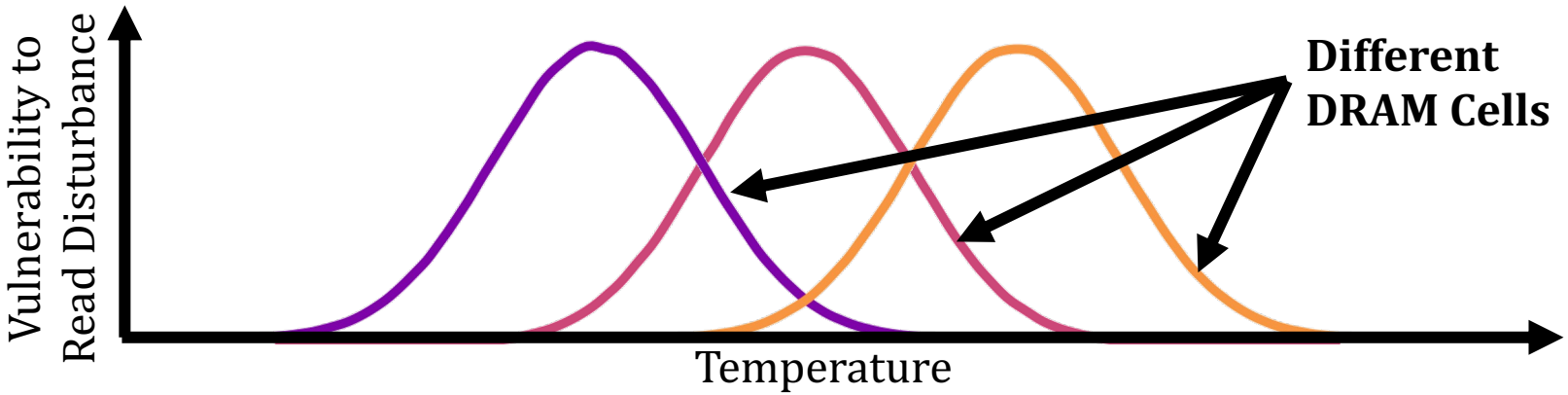
Two DRAM standards

Mfr.	DDR4 DIMMs	DDR3 SODIMMs	# Chips	Density	Die	Org.
A (Micron)	9	1	144 (8)	8Gb (4Gb)	B (P)	x4 (x8)
B (Samsung)	4	1	32 (8)	4Gb (4Gb)	F (Q)	x8 (x8)
C (SK Hynix)	5	1	40 (8)	4Gb (4Gb)	B (B)	x8 (x8)
D (Nanya)	4	-	32 (-)	8Gb (-)	C (-)	x8 (-)

4 Major Manufacturers

272 DRAM Chips in total

Key Findings: Temperature



DRAM read disturbance is more effective **within a bounded temperature range**

Trap-Assisted Charge Leakage Model

- Hammering a wordline **pulls and pushes electrons**
- Electrons **get trapped** and **exacerbate charge leakage**, leading to cause bitflips
- With **increasing temperature**, it becomes **less likely for an electron to get trapped**

Vulnerable temperature range varies **across cells**

A DRAM cell should be tested
at **each possible** operating temperature

Contributions to Understanding RowHammer

- Lois Orosa*, **Abdullah Giray Yağlıkçı***, Haocong Luo, Ataberk Olgun, Jisung Park, Hasan Hassan, Minesh Patel, Jeremie S. Kim, and Onur Mutlu,
  "A Deeper Look into RowHammer's Sensitivities: Experimental Analysis of Real DRAM Chips and Implications on Future Attacks and Defenses"
  *Proceedings of the 54th International Symposium on Microarchitecture (**MICRO**)*, Virtual, October 2021.
  [Slides (pptx) (pdf)] [Talk Video (21 minutes)]
  [Short Talk Slides (pptx) (pdf)]
  [Lightning Talk Slides (pptx) (pdf)] [Lightning Talk Video (1.5 minutes)]
  [arXiv version]

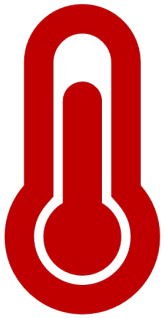
A Follow-Up Work: SpyHammer [Orosa+, arXiv'22]

SpyHammer: Using RowHammer to Remotely Spy on Temperature

Lois Orosa[1,2] Ulrich Rührmair[3,4] A. Giray Yağlıkçı[1] Haocong Luo[1] Ataberk Olgun[1]
Patrick Jattke[1] Minesh Patel[1] Jeremie Kim[1] Kaveh Razavi[1] Onur Mutlu[1]
[1]*ETH Zürich* [2]*Galicia Supercomputing Center (CESGA)* [3]*LMU München* [4]*University of Connecticut*

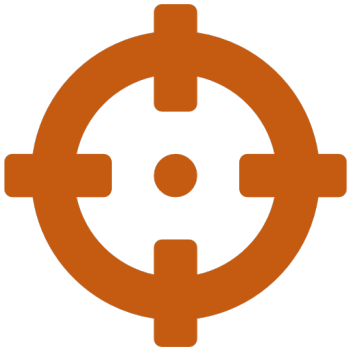
https://arxiv.org/pdf/2210.04084.pdf

Our Recent Works



Temperature

Memory access patterns

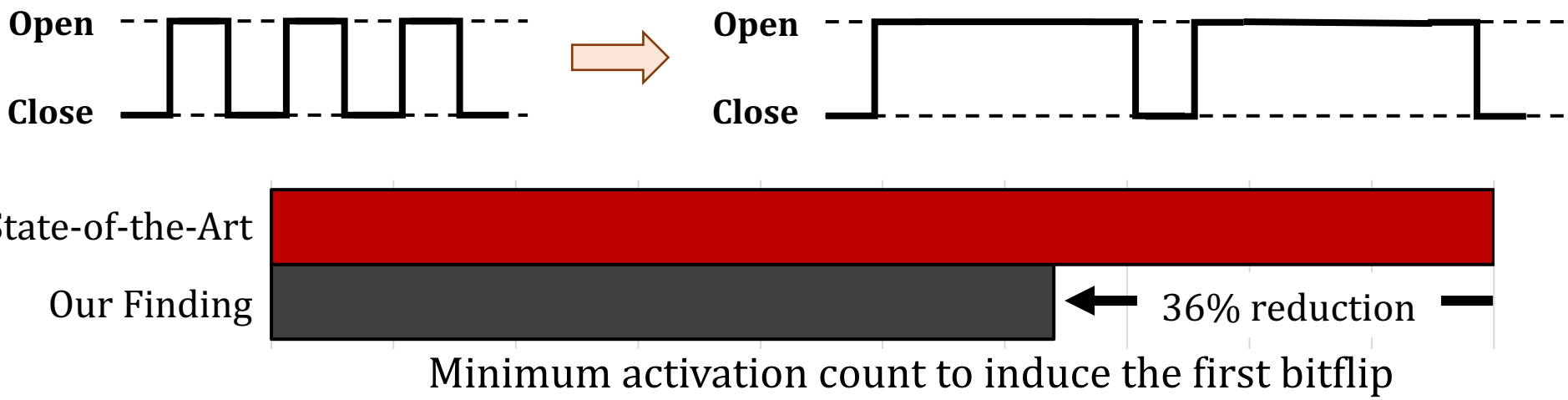


Victim cell's physical location

Leveraging Heterogeneity

Key Findings: Memory Access Patterns

Read disturbance is **more effective**
if the **activated aggressor row** stays **active longer**

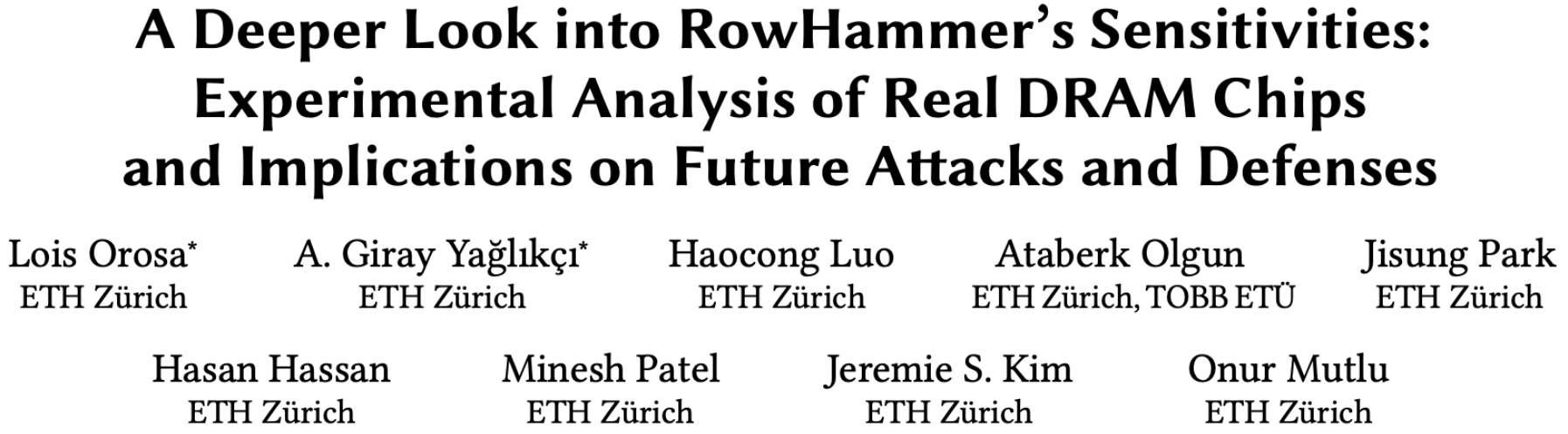


Fewer reads cause a **more significant** read disturbance
when the activated aggressor row stays **active longer**

Existing mitigations are **ineffective** without this insight

Contributions to Understanding RowHammer

- Lois Orosa*, **Abdullah Giray Yağlıkçı***, Haocong Luo, Ataberk Olgun, Jisung Park, Hasan Hassan, Minesh Patel, Jeremie S. Kim, and Onur Mutlu,
  "A Deeper Look into RowHammer's Sensitivities: Experimental Analysis of Real DRAM Chips and Implications on Future Attacks and Defenses"
  *Proceedings of the 54th International Symposium on Microarchitecture (**MICRO**)*, Virtual, October 2021.
  [Slides (pptx) (pdf)] [Talk Video (21 minutes)]
  [Short Talk Slides (pptx) (pdf)]
  [Lightning Talk Slides (pptx) (pdf)] [Lightning Talk Video (1.5 minutes)]
  [arXiv version]



RowPress [Luo+, ISCA 2023] (Follow up of our analysis)

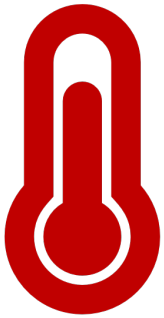
- Haocong Luo, Ataberk Olgun, Giray Yaglikci, Yahya Can Tugrul, Steve Rhyner, M. Banu Cavlak, Joel Lindegger, Mohammad Sadrosadati, and Onur Mutlu,
  "RowPress: Amplifying Read Disturbance in Modern DRAM Chips"
  Proceedings of the 50th International Symposium on Computer Architecture (***ISCA***), Orlando, FL, USA, June 2023.
  [Slides (pptx) (pdf)] [Lightning Talk Slides (pptx) (pdf)] [Lightning Talk Video (3 min)]
  [RowPress Source Code and Datasets (Officially Artifact Evaluated with All Badges)]
  Best artifact award at ISCA 2023.

Keep rows active for **36ns: 47K activations** are enough to induce bitflips
Keep rows active for **7.8us: 5K activations** are enough to induce bitflips
Keep rows active for **30ms: 1 activation** is enough to induce bitflips

RowPress: Amplifying Read-Disturbance in Modern DRAM Chips

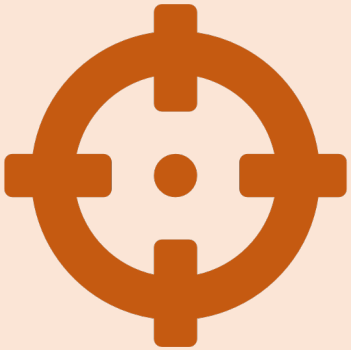
Haocong Luo Ataberk Olgun A. Giray Yağlıkçı Yahya Can Tuğrul Steve Rhyner
Meryem Banu Cavlak Joël Lindegger Mohammad Sadrosadati Onur Mutlu
ETH Zürich

Our Recent Works



Temperature

Memory access patterns

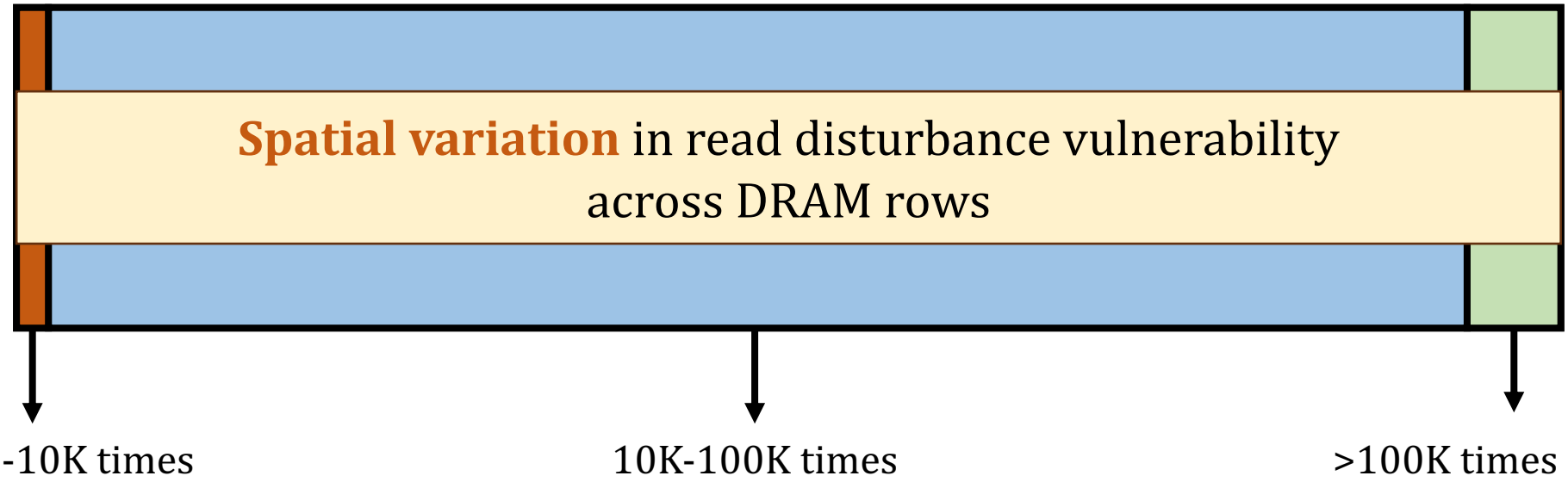


Victim cell's physical location

Leveraging Heterogeneity

Variation in Read Disturbance Vulnerability Across DRAM Rows

- To induce a **read disturbance bitflip,** one should access a row



- Read disturbance solutions are configured **for the worst row**
- **Not all rows** need the **same level** of protection
- Read disturbance solutions incur **large performance overheads** due to **overprotecting many rows**

Tested DRAM Chips

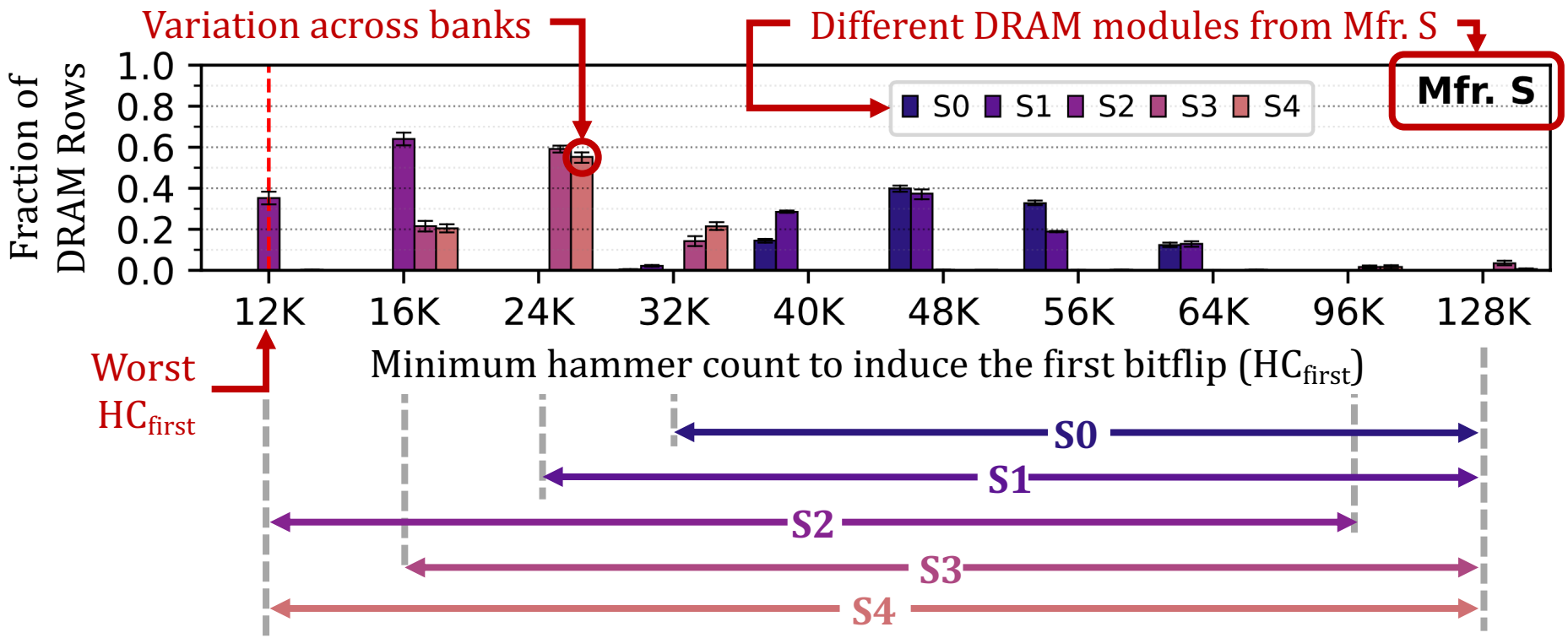
144 DRAM chips from SK Hynix, Micron, and Samsung

Mfr.	DIMM ID	# of Chips	Density Die Rev.	Chip Org.	Date (ww-yy)
Mfr. H (SK Hynix)	H0	8	16Gb – A	x8	51-20
	H1, H2, H3	3 × 8	16Gb – C	x8	48-20
	H4	8	8Gb – D	x8	48-20
Mfr. M (Micron)	M0	4	16Gb – E	x16	46-20
	M1, M3	2 × 16	8Gb – B	x4	N/A
	M2	16	16Gb – E	x4	14-20
	M4	4	16Gb – B	x16	26-21
Mfr. S (Samsung)	S0, S1	2 × 8	8Gb – B	x8	52-20
	S2	8	8Gb – B	x8	10-21
	S3	8	4Gb – F	x8	N/A
	S4	16	8Gb – C	x4	35-21

Key Takeaway from Real Chip Experiments

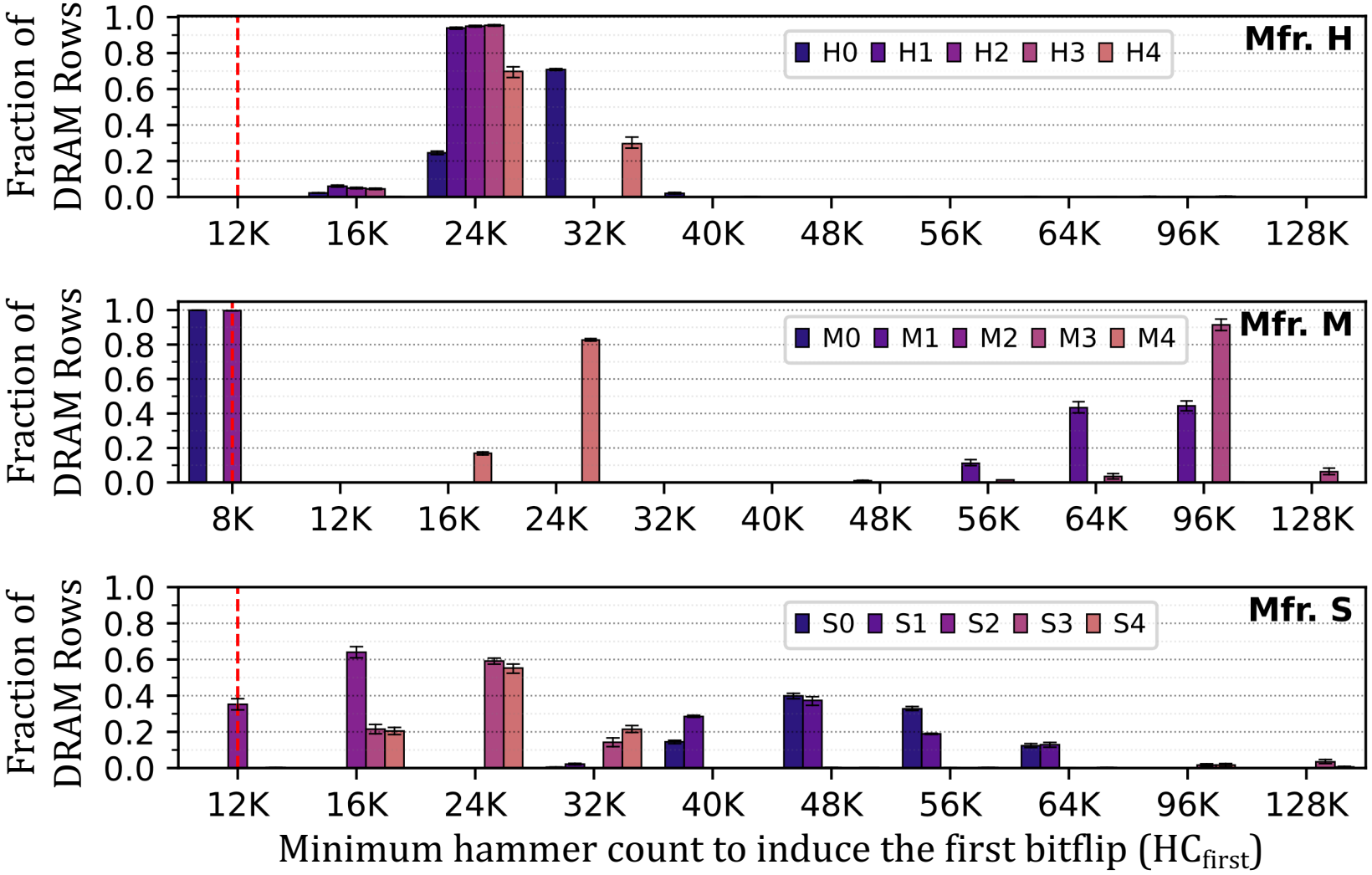
Read disturbance vulnerability varies
significantly and **irregularly**
across DRAM rows

Spatial Variation in the Minimum Hammer Count to Induce the First Bitflip across DRAM Rows

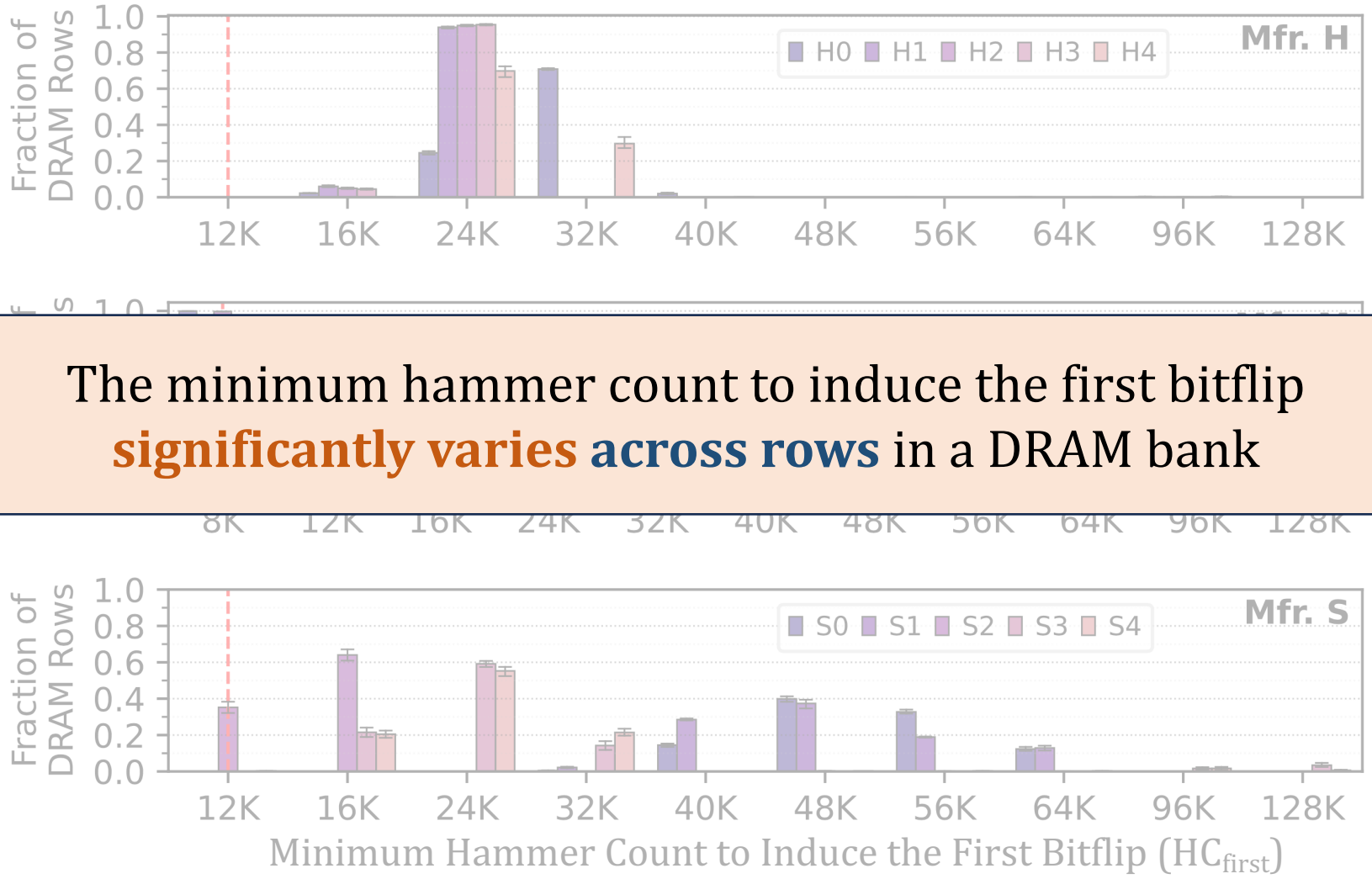


The minimum hammer count to induce the first bitflip **significantly varies across rows** in a DRAM bank

Spatial Variation in the Minimum Hammer Count to Induce the First Bitflip across DRAM Rows



Spatial Variation in the Minimum Hammer Count to Induce the First Bitflip across DRAM Rows



The minimum hammer count to induce the first bitflip **significantly varies across rows** in a DRAM bank

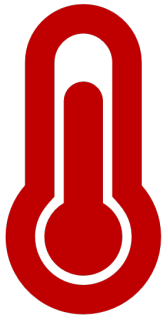
More Detailed Information in the Extended Version

Table 5: Characteristics of the tested DDR4 DRAM modules.

Module Label	Module Vendor	Module Identifier / Chip Identifier	Freq (MT/s)	Mfr. Date ww-yy	Chip Den.	Die Rev.	Chip Org.	Num. of Rows per Bank	HC_{first} Min.	HC_{first} Avg.	HC_{first} Max.
H0	SK Hynix	HMAA4GU6AJR8N-XN [287] H5ANAG8NAJR-XN [288]	3200	51-20	16Gb	A	×8	128K	16K	46.2K	96K
H1		HMAA4GU7CJR8N-XN [289] H5ANAG8NCJR-XN [231]	3200	51-20	16Gb	C	×8	128K	12K	54.0K	128K
H2		HMAA4GU7CJR8N-XN [289] H5ANAG8NCJR-XN [231]	3200	36-21	16Gb	C	×8	128K	12K	55.4K	128K
H3		HMAA4GU7CJR8N-XN [289] H5ANAG8NCJR-XN [231]	3200	36-21	16Gb	C	×8	128K	12K	57.8K	128K
H4		KSM32RD8/16HDR [290] H5AN8G8NDJR-XNC [232]	3200	48-20	8Gb	D	×8	64K	16K	38.1K	96K
M0	Micron	MTA4ATF1G64HZ-3G2E1 [233] MT40A1G16KD-062E [234]	3200	46-20	16Gb	E	×16	128K	8K	24.5K	40K
M1		MTA18ASF2G72PZ-2G3B1QK [235] MT40A2G4WE-083E:B [291]	2400	N/A	8Gb	B	×4	128K	40K	64.5K	96K
M2		MTA36ASF8G72PZ-2G9E1TI [236] MT40A4G4JC-062E:E [292]	2933	14-20	16Gb	E	×4	128K	8K	28.6K	48K
M3		MTA18ASF2G72PZ-2G3B1QK [235] MT40A2G4WE-083E:B [291]	2400	36-21	8Gb	B	×4	128K	56K	90.0K	128K
M4		MTA4ATF1G64HZ-3G2B2 [237] MT40A1G16RC-062E:B [293]	3200	26-21	16Gb	B	×16	128K	12K	42.2K	96K
S0	Samsung	M393A1K43BB1-CTD [294] K4A8G085WB-BCTD [230]	2666	52-20	8Gb	B	×8	64K	32K	57.0K	128K
S1		M393A1K43BB1-CTD [294] K4A8G085WB-BCTD [230]	2666	52-20	8Gb	B	×8	64K	24K	59.8K	128K
S2		M393A1K43BB1-CTD [294] K4A8G085WB-BCTD [230]	2666	10-21	8Gb	B	×8	64K	12K	42.7K	96K
S3		F4-2400C17S-8GNT [295] K4A4G085WF-BCTD [296]	2400	04-21	4Gb	F	×8	32K	16K	59.2K	128K
S4		M393A2K40CB2-CTD [229] K4A8G045WC-BCTD [297]	2666	35-21	8Gb	C	×4	128K	12K	55.4K	128K

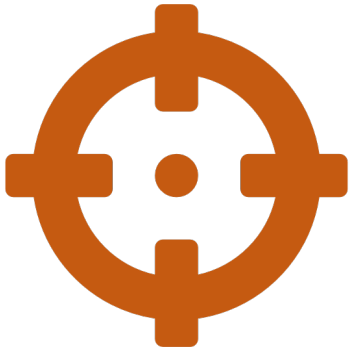
https://arxiv.org/pdf/2402.18652.pdf

Our Recent Works



Temperature

Memory access patterns



Victim cell's physical location

Leveraging Heterogeneity

Motivation

DRAM read disturbance worsens
as DRAM **chip density increases**

Existing solutions become **more aggressive**

Overprotecting many rows significantly increases their **performance overhead**

Motivation

Can we **leverage the variation** in read disturbance **vulnerability** across **DRAM rows**

to **reduce the performance overheads** of existing read disturbance solutions?

Svärd: Spatial Variation-Aware Read Disturbance Defenses

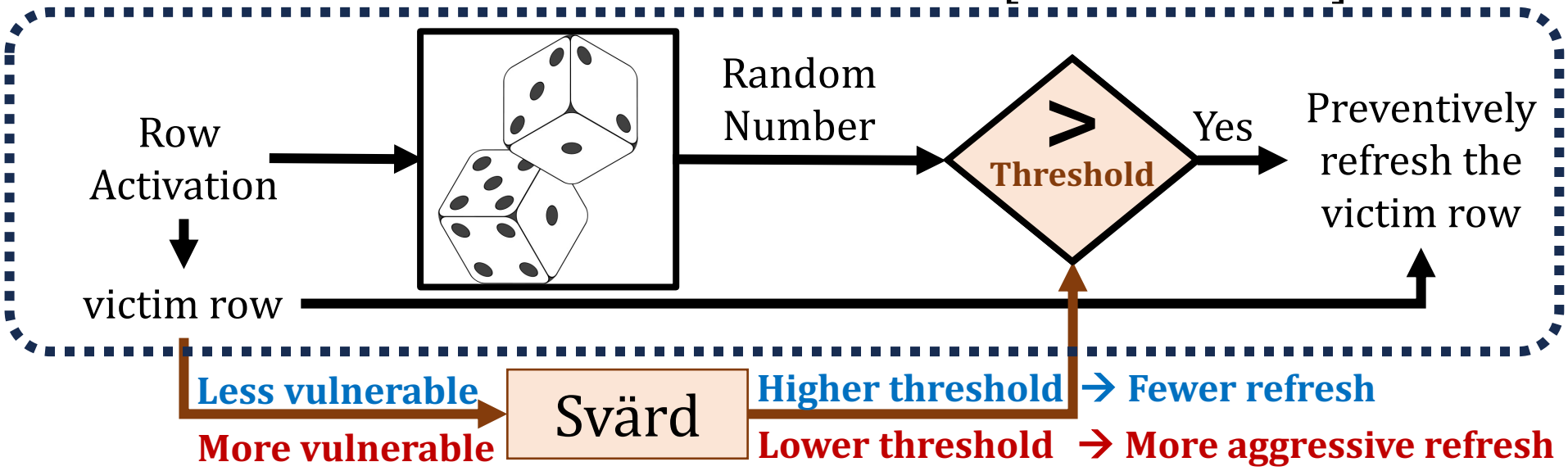
- **Dynamically tunes the aggressiveness** of existing solutions to **the victim row's read disturbance vulnerability**
- Svärd performs **fewer preventive actions (e.g., refresh)** for rows that are **less vulnerable to read disturbance**
- Svärd performs **more preventive actions (e.g., refresh)** for rows that are **more vulnerable to read disturbance**

Svärd significantly **reduces the performance overhead** of existing solutions

Svärd is Swedish for sword, a weapon with a long blade for cutting or thrusting [Merriam-Webster]

Svärd: Integration with Existing Read Disturbance Solutions

PARA: Probabilistic Row Activation [Kim+, ISCA'14]



Svärd dynamically tunes **PARA's threshold** to the **victim row's vulnerability**

Svärd works with **many read disturbance solutions**, including:

BlockHammer [Yaglikci+, HPCA'21]

Hydra [Qureshi+, ISCA'22]

RRS [Saileshwar+, ASPLOS'22]

AQUA [Saxena+, MICRO'22]

Svärd: Metadata Management

- Classifies DRAM rows into **several vulnerability bins**
  Maintains **a few (e.g., four) bits** per DRAM row
- Implemented **where the read disturbance solution is**
- Memory controller-based implementation:
  Metadata can be maintained in
  - SRAM table in the memory controller
  - Data integrity bits in the DRAM chip
  - …
- In-DRAM implementation:
  Metadata can be maintained in
  - DRAM rows
  - Separate DRAM array
  - …

Performance Evaluation

- Cycle-level simulations using **Ramulator 2.0** [Luo+, CAL 2023]
  https://github.com/CMU-SAFARI/ramulator2

- **System Configuration**:

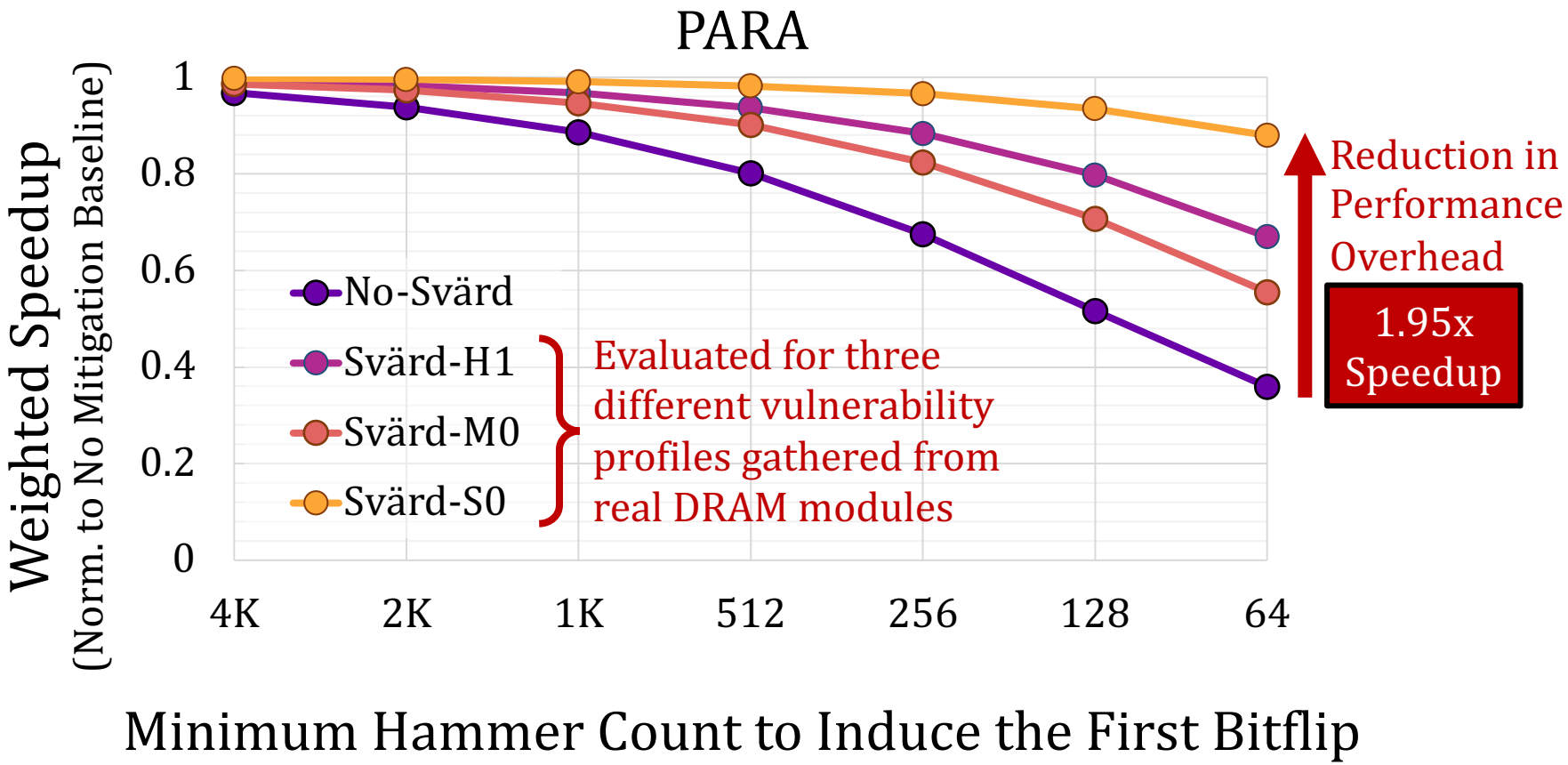
Processor	3.2 GHz, 8 core, 4-wide issue, 128-entry instr. window
Last-Level Cache	64-byte cache line, 8-way set-associative, 8 MB
Memory Scheduler	FR-FCFS
Address Mapping	Minimalistic Open Pages
Main Memory	DDR4, 4 bank group, 4 banks per bank group (16 banks per rank)

- **Workloads**: **120** different **8-core** multiprogrammed workloads from **SPEC CPU2006, SPEC CPU2017, TPC, MediaBench,** and **YCSB** benchmark suites

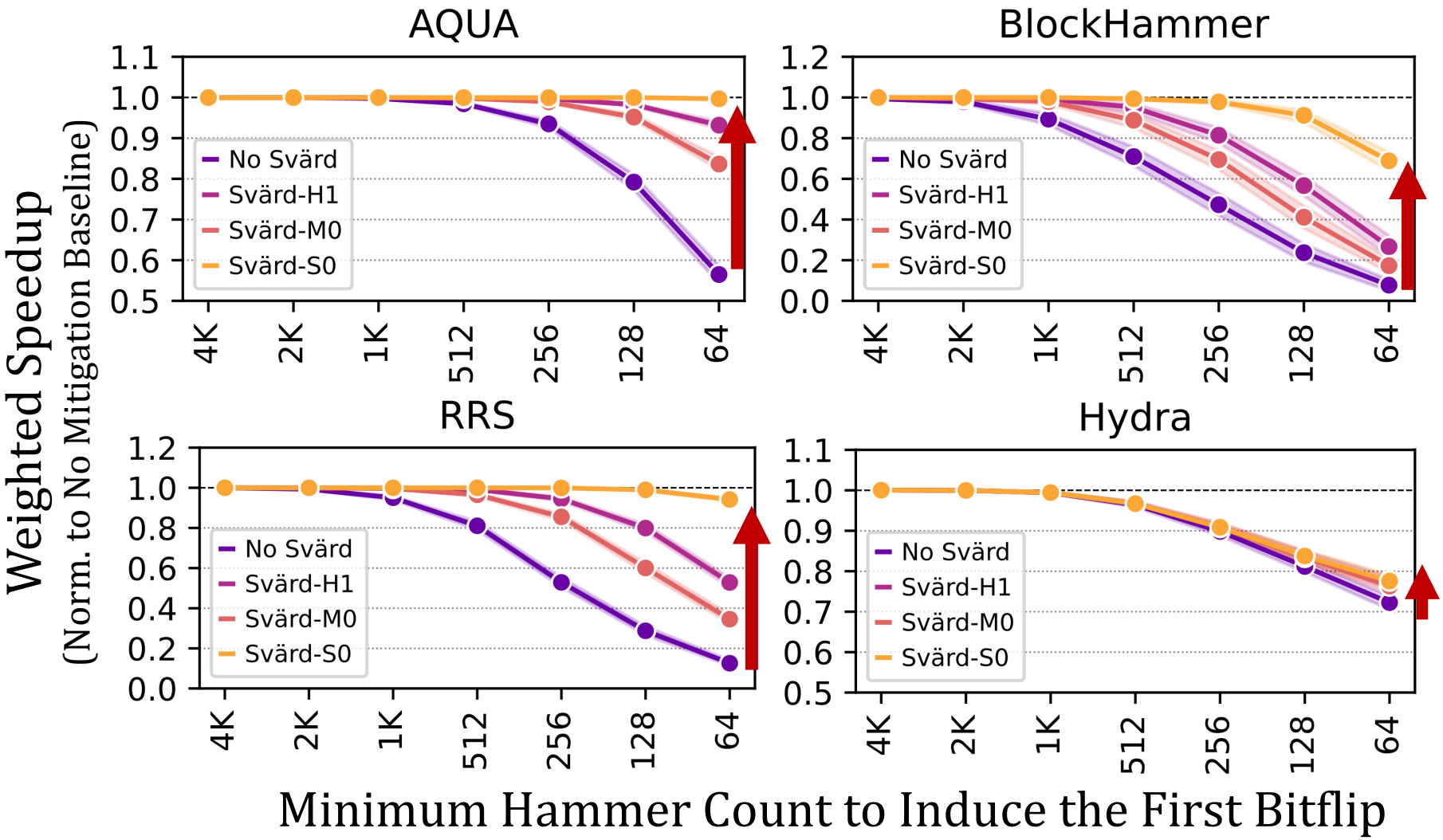
- Integrated with **AQUA, BlockHammer, PARA, Hydra,** and **RRS**

- **HC_{first}**: {4K, 2K, 1K, 512, 256, 128, 64} hammers
  Minimum **hammer count** needed to induce **the first bitflip**

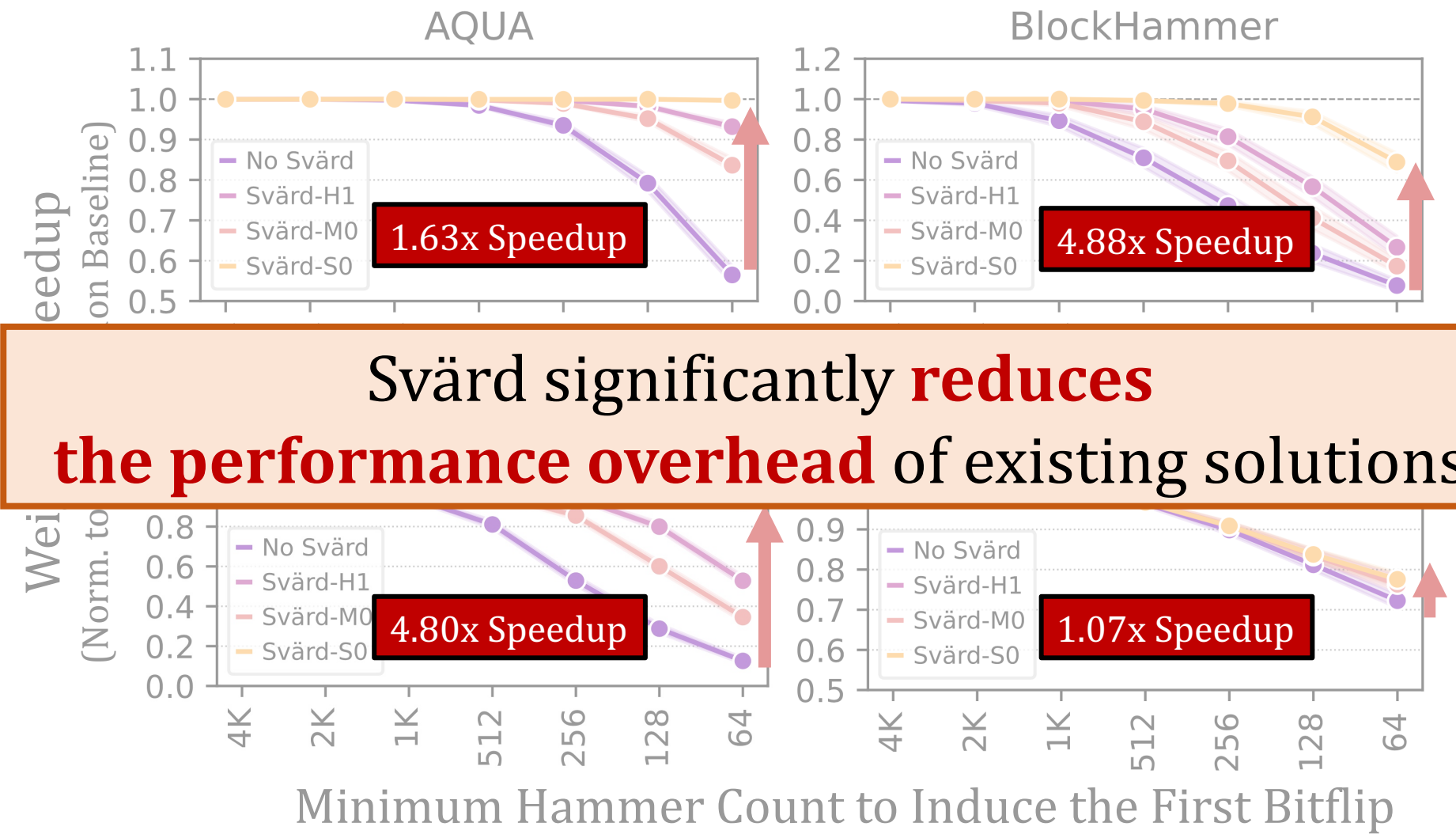
Implications on Future Solutions



Implications on Future Solutions



Implications on Future Solutions



Svärd significantly **reduces**
the performance overhead of existing solutions

Spatial Variation-Aware Read Disturbance Defenses

- **The first rigorous experimental study** on the spatial variation of DRAM read disturbance across DRAM rows
  - **144 DDR4 DRAM chips** from **three major manufacturers**
  - Characterize **all rows in a bank** and four banks in a DRAM chip

Read disturbance vulnerability varies **significantly** and **irregularly** across DRAM rows

- **Svärd: Spatial Variation-Aware Read Disturbance Defenses**
  - **Dynamically tunes** a solution's aggressiveness (e.g., perform more/less refresh) to the **victim row's vulnerability** to DRAM read disturbance
  - Implemented either in **the memory controller** or in **the DRAM chip**

Svärd significantly **reduces** **the performance overhead** of existing solutions

Svärd may present itself to any worthy read disturbance solution

Enabling Efficient and Scalable Solutions

Spatial Variation-Aware Read Disturbance Defenses: Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

Abdullah Giray Yağlıkçı Geraldo F. Oliveira Yahya Can Tuğrul
İsmail Emir Yüksel Ataberk Olgun Haocong Luo Onur Mutlu
ETH Zürich

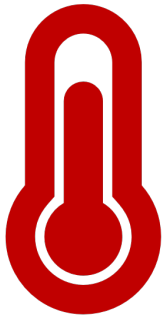
Read disturbance in modern DRAM chips is a widespread phenomenon and is reliably used for breaking memory isolation, a fundamental building block for building robust systems. RowHammer and RowPress are two examples of read disturbance in DRAM where repeatedly accessing (hammering) or keeping active (pressing) a memory location induces bitflips in other memory locations. Unfortunately, shrinking technology node size exacerbates read disturbance in DRAM chips over generations. As a result, existing defense mechanisms suffer from significant performance and energy overheads, limited effectiveness, or prohibitively high hardware complexity.

In this paper, we tackle these shortcomings by leveraging the spatial variation in read disturbance across different memory locations in real DRAM chips. To do so, we 1) present the

Many prior works demonstrate attacks on a wide range of systems that exploit read disturbance to escalate privilege, leak private data, and manipulate critical application outputs [1, 3–53, 71–84]. To make matters worse, various experimental studies [1, 1, 25, 33, 36, 37, 61, 70] find that newer DRAM chip generations are more susceptible to read disturbance. For example, chips manufactured in 2018-2020 can experience RowHammer bitflips at an order of magnitude fewer row activations compared to the chips manufactured in 2012-2013 [61]. As read disturbance in DRAM chips worsens, ensuring robust (i.e., reliable, secure, and safe) operation becomes more expensive in terms of performance overhead, energy consumption, and hardware complexity [61, 85, 86]. Therefore, it is critical to understand the read disturbance vulnerabilities

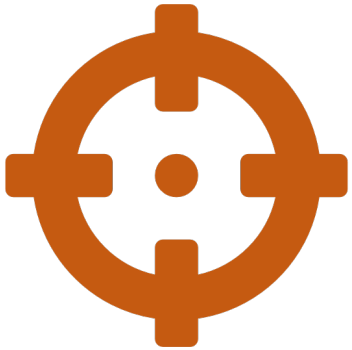
https://arxiv.org/pdf/2402.18652.pdf

Our Recent Works



Temperature

Memory access patterns

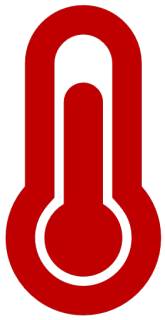


Victim cell's physical location

Leveraging Heterogeneity

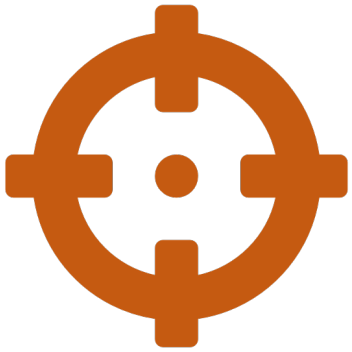
My Dissertation Works

- A deeper look into DRAM read disturbance



Temperature

Memory access patterns



Victim cell's physical location

Voltage

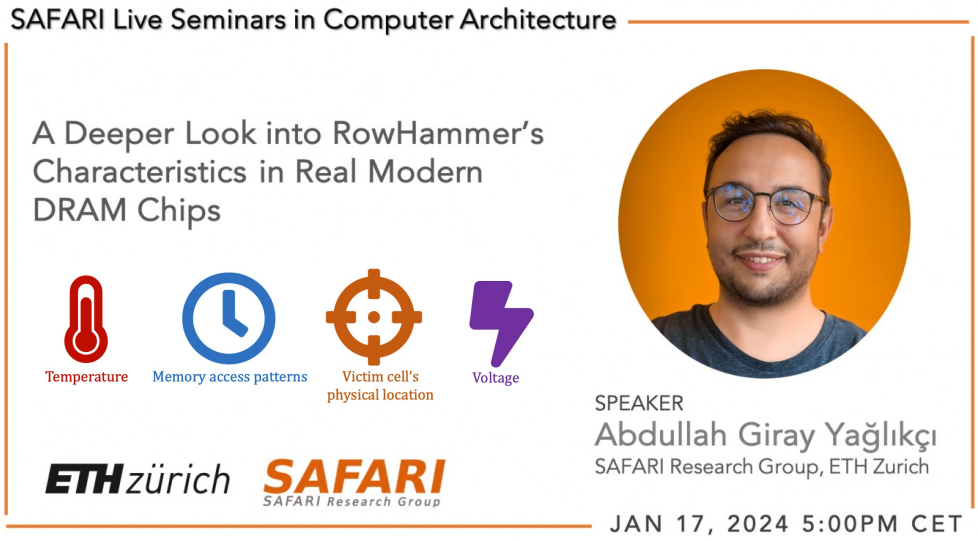
- Solutions to DRAM read disturbance

Leveraging Heterogeneity

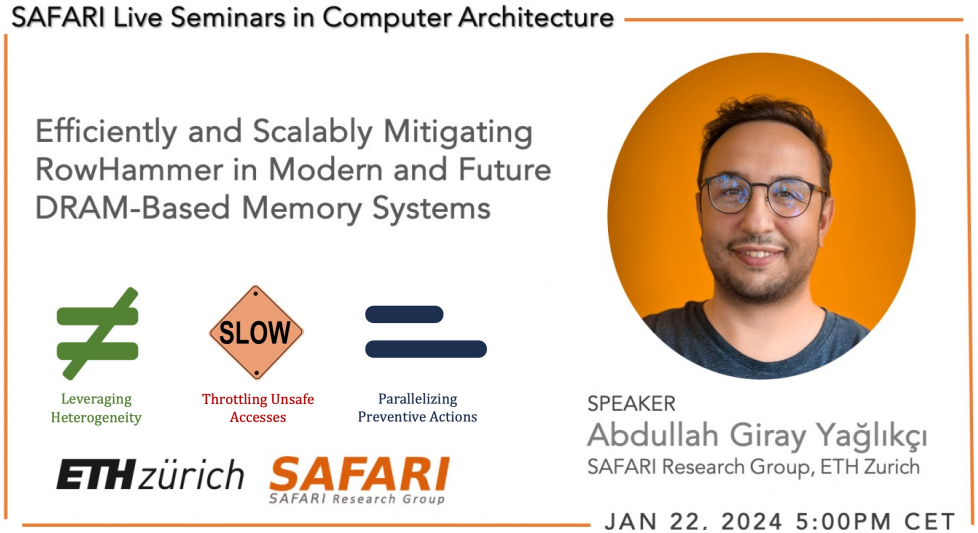
Throttling Unsafe Accesses

Parallelizing Preventive Actions

More Details and Discussion on YouTube



https://www.youtube.com/live/CRtm1es4n3o?si=8N5zB6e_RUc5Ejl8



https://www.youtube.com/live/YQwRYWpCsk0?si=jXPueMHb5wgs69-q

Enabling Efficient and Scalable Read Disturbance Mitigation

via New Experimental Insights into Modern Memory Chips

agyaglikci.github.io

Abdullah Giray Yaglikci

agyaglikci@gmail.com

https://agyaglikci.github.io

13 March 2024

ARM Cambridge

safari.ethz.ch

SAFARI

ETH zürich

Enabling Efficient and Scalable Read Disturbance Mitigation via New Experimental Insights into Modern Memory Chips

Current and Future Work Discussion

agyaglikci.github.io

Abdullah Giray Yaglikci

agyaglikci@gmail.com

https://agyaglikci.github.io

11 March 2024

Huawei Cambridge

safari.ethz.ch

SAFARI

ETH zürich

Current and Future Challenges


Reliability


Performance


Fairness


Energy Efficiency

While the memory systems

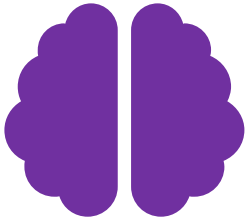
1. **scale-up and are shared** across many users (e.g., disaggregated memory systems)

2. **scale-down** in manufacturing technology node size

3. support **processing near/using memory**

Future Research for Better Memory Systems

Deeper Understanding of
Physics and Vulnerabilities



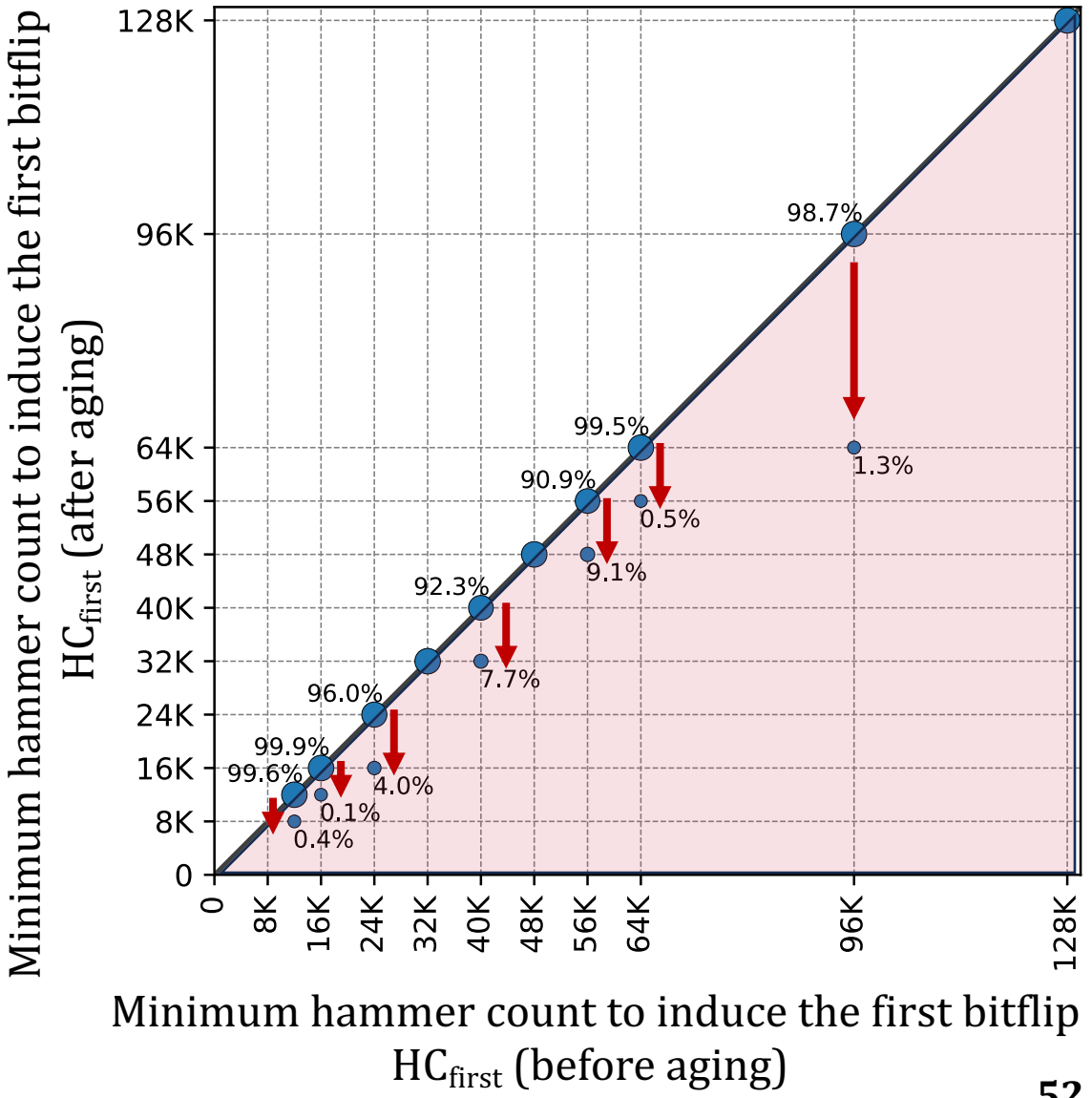
Flexible and Intelligent Memory
Chips, Interfaces, Controllers

Cross-Layer
Communication

Deeper Understanding of Physics and Vulnerabilities

- The effect of **aging**
  Preliminary data on aging via 68-day of continuous hammering

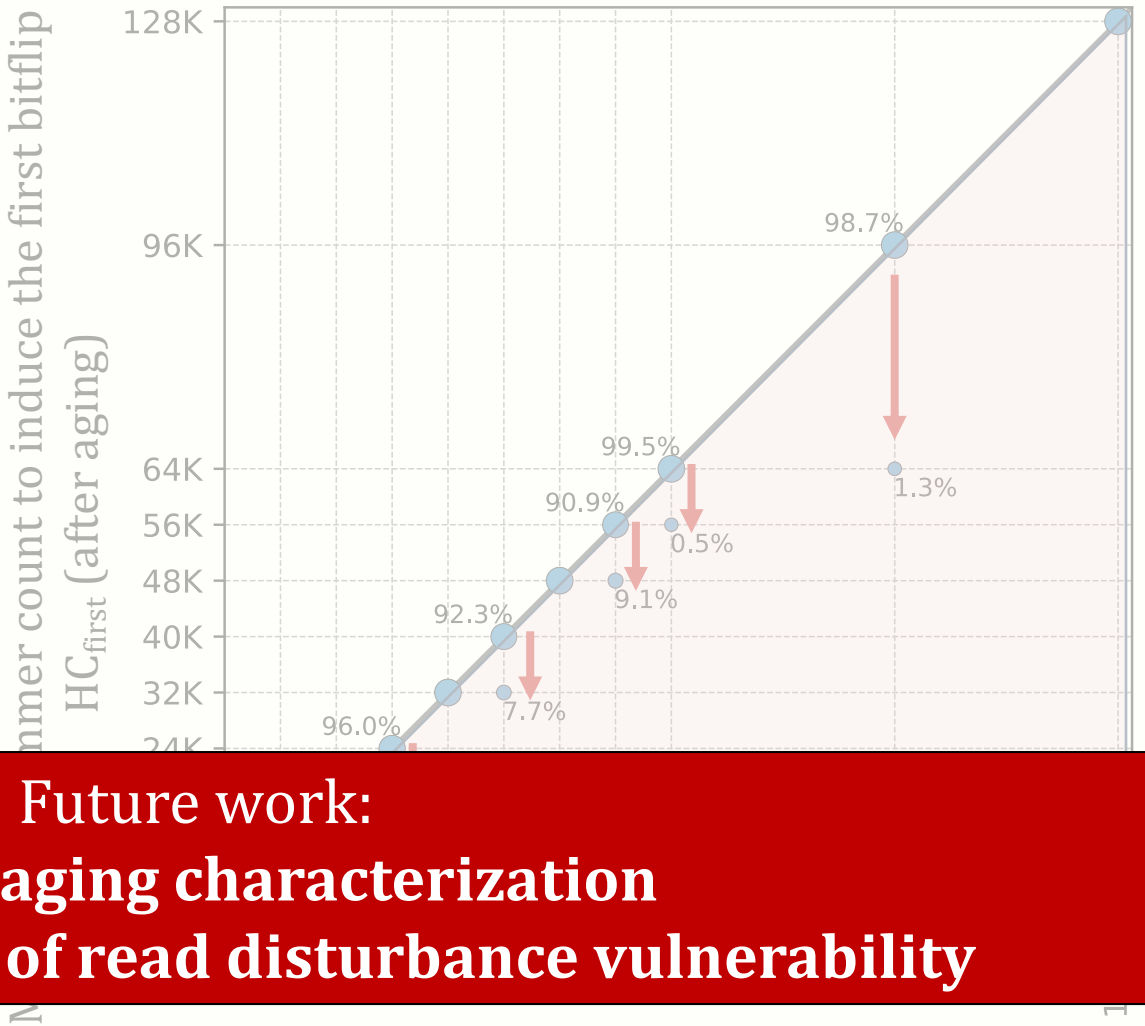
Aging can lead to read disturbance bitflips at **smaller** hammer counts



Deeper Understanding of Physics and Vulnerabilities

- The effect of **aging**
  Preliminary data on aging via 68-day of continuous hammering

Aging can lead to read disturbance bitflips at **smaller** hammer counts



Future work:
rigorous aging characterization
and **online profiling of read disturbance vulnerability**

Deeper Understanding of Physics and Vulnerabilities

- The effect of **aging**

- **Interactions** across different error mechanisms
  - RowHammer
  - RowPress
  - Data retention time errors
  - Variable retention time
  - ...

Deeper Understanding of Physics and Vulnerabilities

- The effect of **aging**
- **Interactions** across different error mechanisms
- What is **the worst-case**?
  - Temperature
  - Data pattern
  - Memory access pattern
  - Spatial variation
  - Voltage

What is **the worst-case** considering all **these sensitivities**?

What is **the minimum hammer count** to induce a read disturbance bitflip?

Deeper Understanding of Physics and Vulnerabilities

- The effect of **aging**
- **Interactions** across different error mechanisms
- What is **the worst-case**?

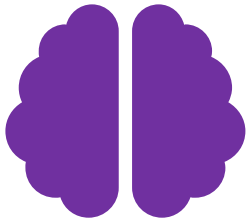
How reliable are our DRAM chips?

How reliable will our DRAM chips be tomorrow?

We **do not** know! This is an **open research problem**

Future Research for Better Memory Systems

Deeper Understanding of
Physics and Vulnerabilities



Flexible and Intelligent Memory
Chips, Interfaces, Controllers

Cross-Layer
Communication

Flexible and Intelligent Chips, Interfaces, Controllers

- **In-field patching** is necessary

DRAM read disturbance solution is here

Configuration

DRAM chip is here

Reconfiguration is needed

+ add more hardware counters
+ increase metadata cache capacity

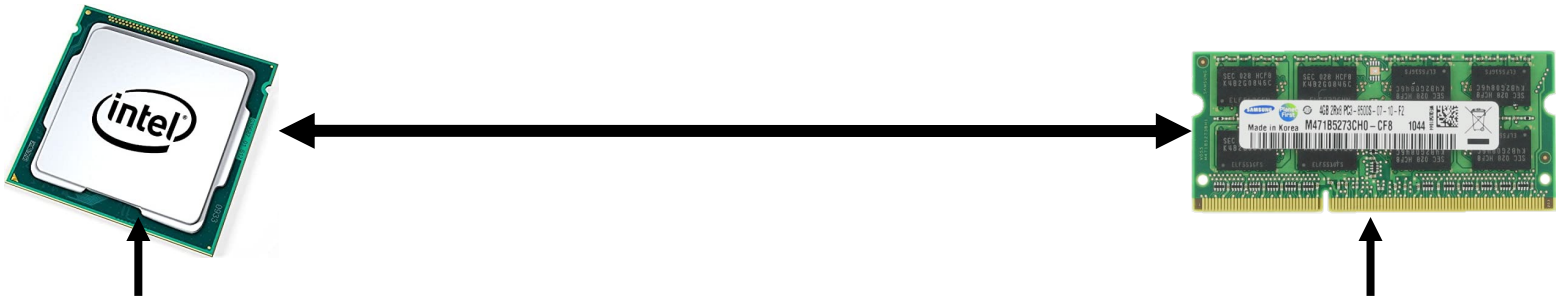
DRAM module can be replaced

Aging might change the chip's characteristics

Deployed solutions should be patchable in field

Flexible and Intelligent Chips, Interfaces, Controllers

- **In-field patching** is necessary
- Interfaces should be **more flexible**



Memory controller decides what should be done when

DRAM chip has read disturbance solution inside (tracking+prevention)

- The memory controller should provide the DRAM chip with **necessary time window** to perform **preventive actions (e.g., refreshing rows)**
- The memory controller **does not have** the tracking information
- Communicating is **not straightforward** due to strict communication protocol

A more flexible interface is necessary

How Large is 1000 Activations?

- Bitflips occur **at ~1000 activations**
- Mitigation mechanisms trigger **preventive actions** (e.g., preventive refresh) **at ~500 activations**
- Is 500 **a distinctive activation count**?
- Benign workloads activate **hundreds of rows** more than **500 times** in a refresh window

Memory intensive workloads

from SPEC'06, SPEC'17, TPC, YCSB, and MediaBench

Workload	MPKI	ACT-64+	ACT-128+	ACT-512+
429.mcf	68.27	2564	2564	2564
470.lbm	28.09	7089	6596	664
462.libquantum	25.95	1	0	0
549.fotonik3d	25.28	10065	88	0
459.GemsFDTD	24.93	10572	218	0
519.lbm	24.37	5824	5455	2482
434.zeusmp	22.24	11085	4825	292
510.parest	17.79	803	185	94
433.milc	17.22	321	92	0
437.leslie3d	15.82	4678	631	7
483.xalancbmk	13.67	4354	776	113
482.sphinx3	12.59	1385	762	304
505.mcf	11.35	1582	1384	732
471.omnetpp	10.72	1015	419	122
tpch2	9.09	875	307	88
520.omnetpp	9.00	1185	84	32
tpch17	7.43	1196	158	26
473.astar	5.18	5957	22	0
436.cactusADM	4.94	6151	2354	1134
jp2_encode	4.18	0	0	0

Benign workloads **might not be so benign**

Flexible and Intelligent Chips, Interfaces, Controllers

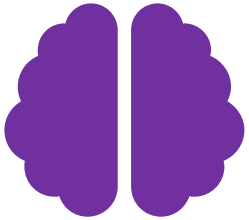
- **In-field patching** is necessary
- Interfaces should be **more flexible**
- Memory controllers should be **more intelligent** in detecting **malicious activity**

- DRAM chips become **more and more vulnerable** to RowHammer **and RowPress**
- **Key Insight:**
  - **A thousand activations** are enough **to induce bitflips**
  - Benign applications **perform as many activations**
- **Problem:** DRAM read disturbance solutions are getting **prohibitively expensive**
- **Research Question:** How to identify malicious threads/processes/users?
- More **intelligent detection** mechanisms are needed → AI can play an important role
- The **memory controller** observes all memory accesses → has the ground truth data

More intelligent memory controllers can help

Future Research for Better Memory Systems

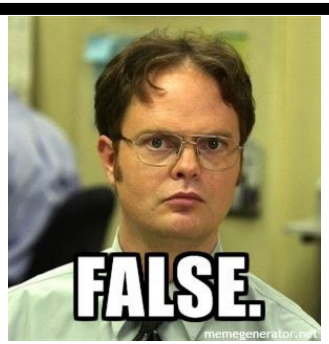
Deeper Understanding of
Physics and Vulnerabilities



Flexible and Intelligent Memory
Chips, Interfaces, Controllers

Cross-Layer
Communication

Cross-Layer Communication

		Detection	Mitigation
Software	• Memory allocations • Memory access patterns • Control flow patterns • Time / power measurements		++
uArch	• Memory request scheduling • Speculative execution • Prefetching, branch prediction • Power management	+	+
Device	• Bitflips occur • Memory isolation is broken	++	~

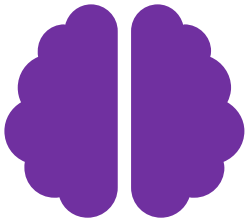
Cross-Layer Communication

		Detection	Mitigation
Software	• Memory allocations • Memory access patterns • Control flow patterns • Time / power measurements		++
Device	• Bitflips occur • Memory isolation is broken	++	~

Cross-layer communication is crucial going forward

Future Research for Better Memory Systems

Deeper Understanding of
Physics and Vulnerabilities



Flexible and Intelligent Memory
Chips, Interfaces, Controllers

Cross-Layer
Communication

Enabling Efficient and Scalable Read Disturbance Mitigation via New Experimental Insights into Modern Memory Chips

Backup Slides

Abdullah Giray Yaglikci

agyaglikci@gmail.com

https://agyaglikci.github.io

agyaglikci.github.io

safari.ethz.ch

SAFARI

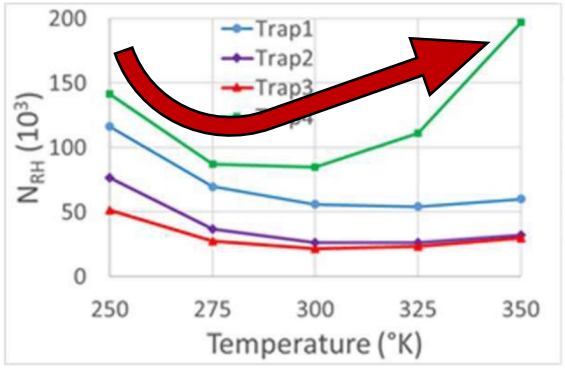
ETH zürich

Circuit-Level Justification

Trap-Assisted Charge Leakage Model

- Hammering a wordline **pulls and pushes electrons**
- Electrons **get trapped** and **exacerbate charge leakage**, leading to cause bitflips
- With **increasing temperature**, it becomes **less likely for an electron to get trapped**

3D TCAD Evaluation [Yang+, EDL'19]



	Row	Column	Energy
Trap1	1	4	0.50eV
Trap2	2	3	0.45eV
Trap3	3	3	0.45eV
Trap4	4	3	0.45eV

Fig. 6. Hammering threshold N_{RH} vs. temperature from 250 to 350°K for different traps. Location in row and column refers to matrix in Fig. 2b.

Until a temperature inflection point:

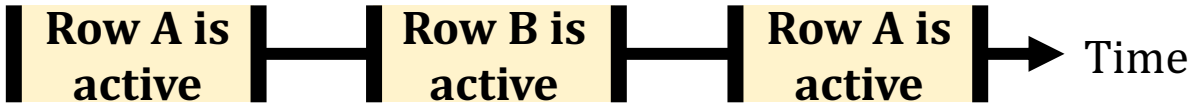
As temperature increases, **fewer activations** can cause bitflips

After the temperature inflection point:

As temperature increases, **more activations** are needed to cause bitflips

Example Attack Improvement: Bypassing Defenses with Aggressor Row Active Time

Activating aggressor rows **as frequently as possible**:



Keeping the aggressor rows **active for a longer time**:

Reduces the minimum activation count to induce a bitflip **by 36%**

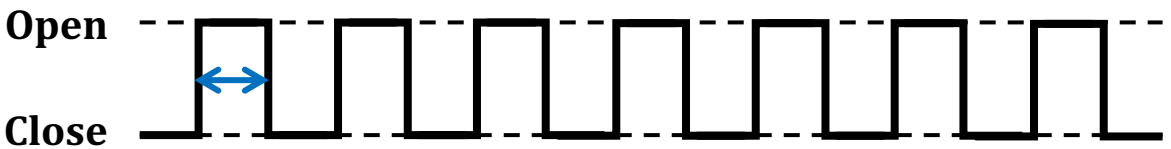
Bypasses defenses that do not account for this reduction

RowPress vs. RowHammer

Instead of using a high activation count,

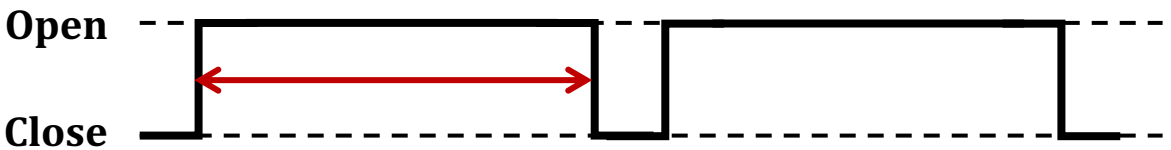
☞ increase the time that the aggressor row stays open

RowHammer Aggressor Row



36ns, 47K activations to induce bitflips

RowPress Aggressor Row



7.8µs, only 5K activations to induce bitflips

We observe bitflips even with **ONLY ONE activation** in extreme cases where the row stays open for 30ms

RowPress [Luo+, ISCA 2023]

- Haocong Luo, Ataberk Olgun, Giray Yaglikci, Yahya Can Tugrul, Steve Rhyner, M. Banu Cavlak, Joel Lindegger, Mohammad Sadrosadati, and Onur Mutlu,
  "RowPress: Amplifying Read Disturbance in Modern DRAM Chips"
  Proceedings of the 50th International Symposium on Computer Architecture (***ISCA***), Orlando, FL, USA, June 2023.
  [Slides (pptx) (pdf)]
  [Lightning Talk Slides (pptx) (pdf)]
  [Lightning Talk Video (3 minutes)]
  [RowPress Source Code and Datasets (Officially Artifact Evaluated with All Badges)]
  Officially artifact evaluated as available, reusable and reproducible.
  Best artifact award at ISCA 2023.

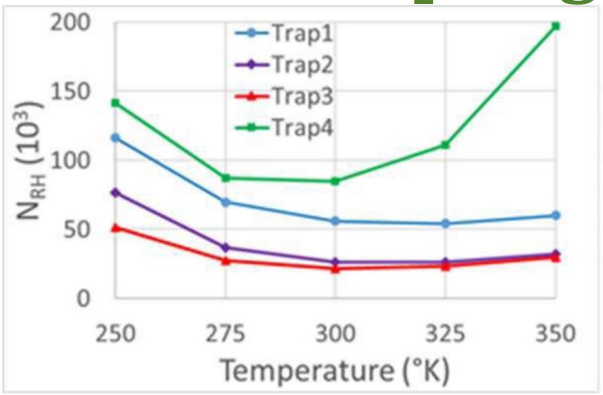
RowPress: Amplifying Read-Disturbance in Modern DRAM Chips

Haocong Luo Ataberk Olgun A. Giray Yağlıkçı Yahya Can Tuğrul Steve Rhyner
Meryem Banu Cavlak Joël Lindegger Mohammad Sadrosadati Onur Mutlu
ETH Zürich

Circuit-Level Justification

We hypothesize that our observations are caused by the **non-monotonic behavior of charge trapping** characteristics of DRAM cells

3D TCAD model [Yang+, EDL'19]



	Row	Column	Energy
Trap1	1	4	0.50eV
Trap2	2	3	0.45eV
Trap3	3	3	0.45eV
Trap4	4	3	0.45eV

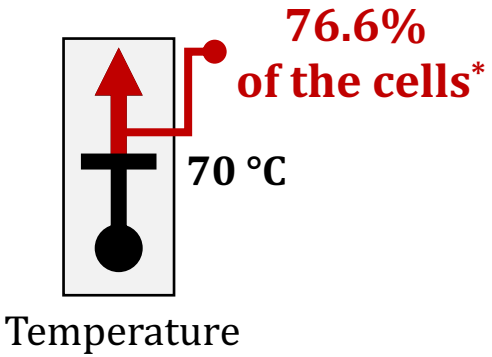
Fig. 6. Hammering threshold N_{RH} vs. temperature from 250 to 350°K for different traps. Location in row and column refers to matrix in Fig. 2b.

***HC_{first}* decreases as temperature increases**, until a temperature inflection point where ***HC_{first}* starts to increase as temperature increases**

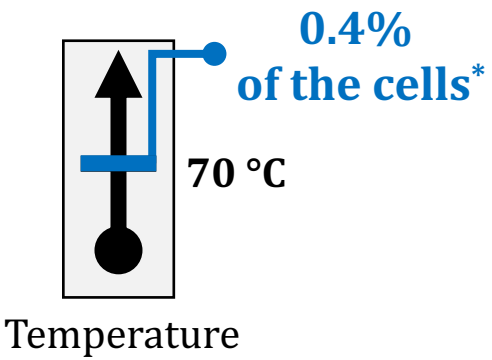
A **cell is more vulnerable** to RowHammer at **temperatures close to its temperature inflection point**

Example Attack Improvement: Temperature-Dependent Trigger

1. Identify **abnormal increase** in temperature to attack a data center **during its peak hours**



2. **Precisely measure** the temperature **to trigger an attack** exactly at the desired temperature



*Example fraction values from SK Hynix modules

Enabling Efficient and Scalable Read Disturbance Mitigation via New Experimental Insights into Modern Memory Chips

Backup Slides

Abdullah Giray Yaglikci

agyaglikci@gmail.com

https://agyaglikci.github.io

11 March 2024

Huawei Cambridge

agyaglikci.github.io

safari.ethz.ch

SAFARI

ETH zürich

Thesis Statement

Developing

a deeper understanding of DRAM read disturbance

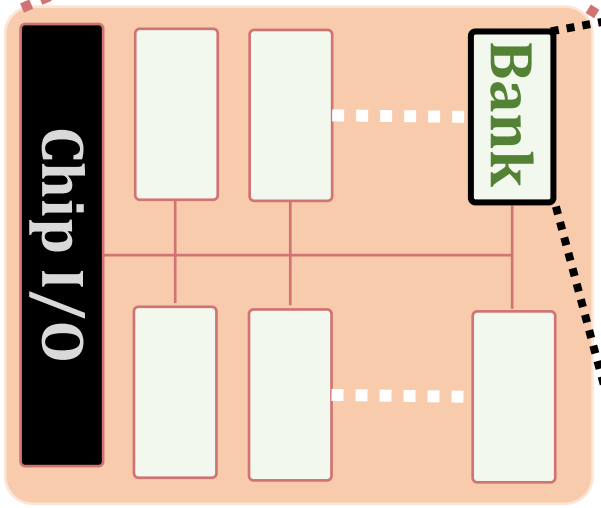
and

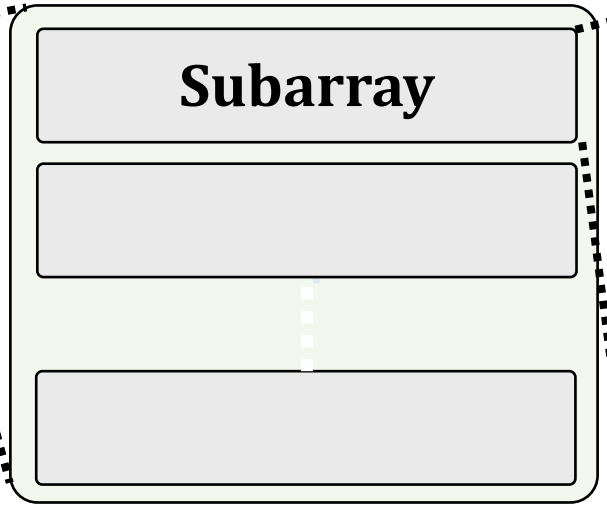
revisiting memory controller designs

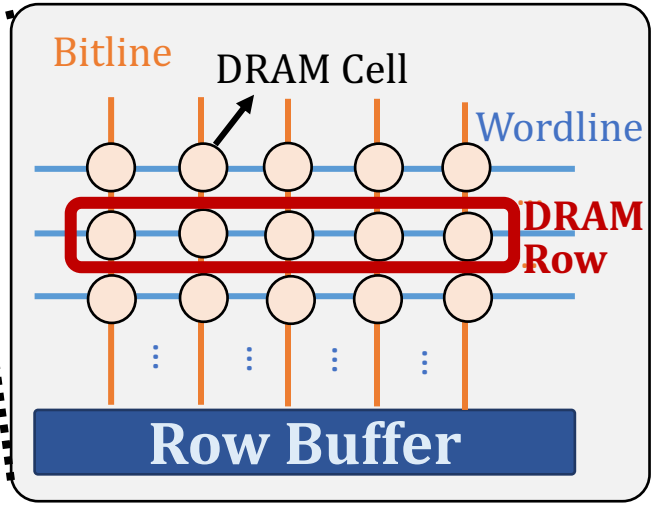
enable scientists and engineers to build

reliable, secure, and safe DRAM-based systems

DRAM Organization







DRAM Chip **DRAM Bank** **DRAM Subarray**

DRAM Chips Tested

A Deeper Look into RowHammer's Sensitivities: Experimental Analysis of Real DRAM Chips and Implications on Future Attacks and Defenses

Lois Orosa* ETH Zürich
A. Giray Yağlıkçı* ETH Zürich
Haocong Luo ETH Zürich
Ataberk Olgun ETH Zürich, TOBB ETÜ
Jisung Park ETH Zürich
Hasan Hassan ETH Zürich
Minesh Patel ETH Zürich
Jeremie S. Kim ETH Zürich
Onur Mutlu ETH Zürich

Table 4: Characteristics of the tested DDR4 and DDR3 DRAM modules.

Type	Chip Manufacturer	Chip Identifier	Module Vendor	Module Identifier	Freq. (MT/s)	Date Code	Density	Die Rev.	Org.	#Modules	#Chips
DDR4	A: Micron	MT40A2G4WE-083E:B	Micron	MTA18ASF2G72PZ-2G3B1QG [94]	2400	1911	8Gb	B	x4	6	96
						1843				2	32
						1844				1	16
	B: Samsung	K4A4G085WF-BCTD [132]	G.SKILL	F4-2400C17S-8GNT [35]	2400	2021 Jan ★	4Gb	F	x8	4	32
	C: SK Hynix	DWCW (Partial Marking) †	G.SKILL	F4-2400C17S-8GNT [35]	2400	2042	4Gb	B	x8	5	40
	D: Nanya	D1028AN9CPGRK ‡	Kingston	KVR24N17S8/8 [75]	2400	2046	8Gb	C	x8	4	32
DDR3	A: Micron	MT41K512M8DA-107:P [22]	Crucial	CT51264BF160BJ.M8FP	1600	1703	4Gb	P	x8	1	8
	B: Samsung	K4B4G0846Q	Samsung	M471B5173QH0-YK0 [131]	1600	1416	4Gb	Q	x8	1	8
	C: SK Hynix	H5TC4G83BFR-PBA	SK Hynix	HMT451S6BFR8A-PB [139]	1600	1535	4Gb	B	x8	1	8

SAFARI

DRAM Testing Methodology

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

1. Prevent sources of interference during core test loop

- **No DRAM refresh**: to avoid refreshing victim row
- **No DRAM calibration events**: to minimize variation in test timing
- **No RowHammer mitigation mechanisms**: to observe circuit-level effects
- Test for **less than a refresh window (32ms)** to avoid retention failures
- **Repeat tests** for ten times

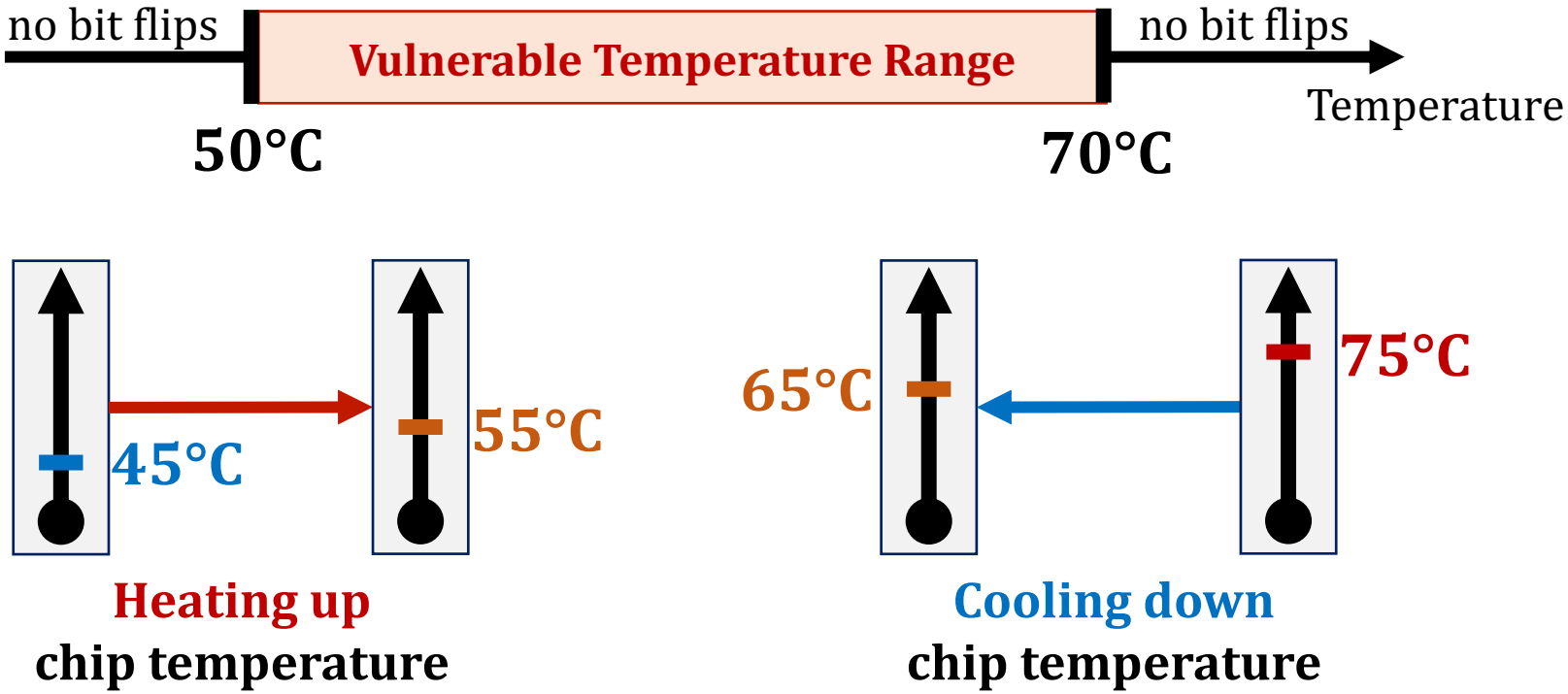
2. Worst-case access sequence

- We use **worst-case** access sequence based on prior works' observations
- For each row, **repeatedly access the two physically-adjacent rows as fast as possible**

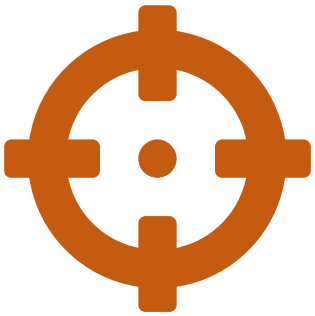
Attack Improvement 1:
Making DRAM Cells More Vulnerable

An attacker can **manipulate temperature** to make the cells that store sensitive data **more vulnerable**

DRAM cells are vulnerable in a **bounded temperature range**



Key Takeaways from Spatial Variation Analysis



In-Chip Variations

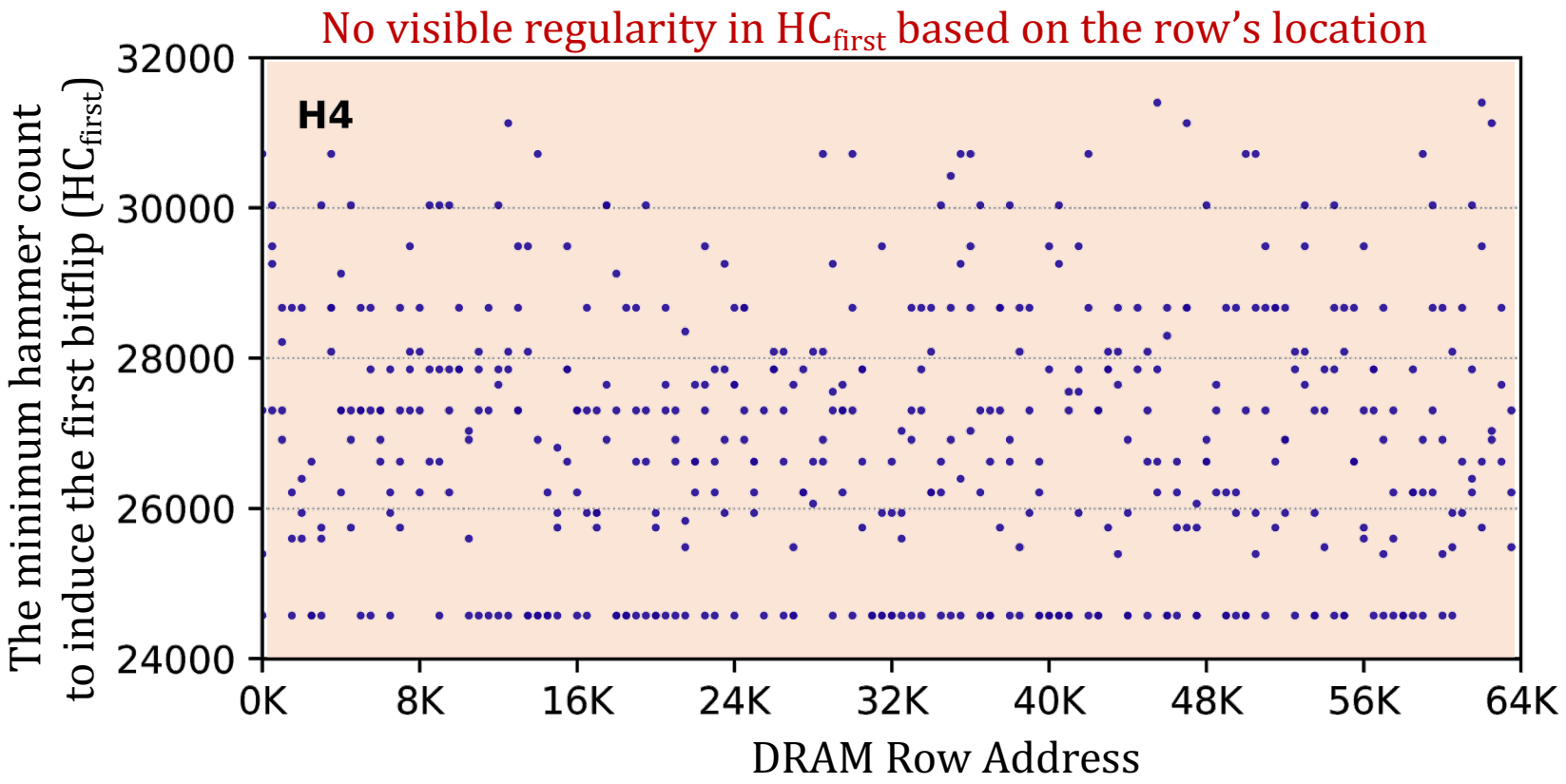
Key Takeaway 1

RowHammer vulnerability **significantly varies** across DRAM rows and columns due to **design** and **manufacturing-process**

Key Takeaway 2

Minimum hammer count to induce the first bitflip (HC_{first}) significantly varies **across rows in a subarray** but **not as much across subarrays**

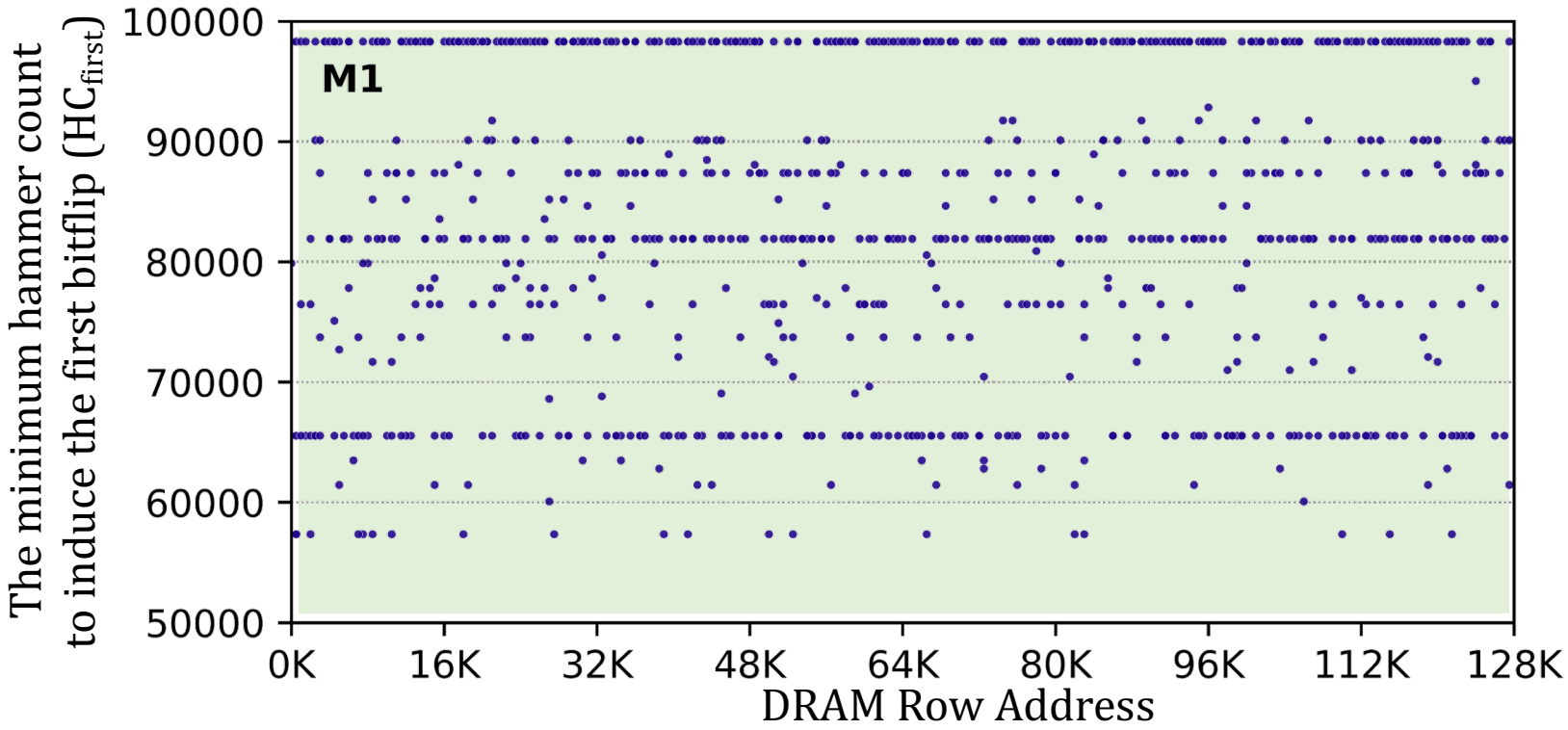
Regularity in Spatial Variation of Read Disturbance across DRAM Rows



The minimum hammer count to induce the first bitflip **irregularly varies** with respect to row's location in DRAM bank

Regularity in Spatial Variation of Read Disturbance across DRAM Rows

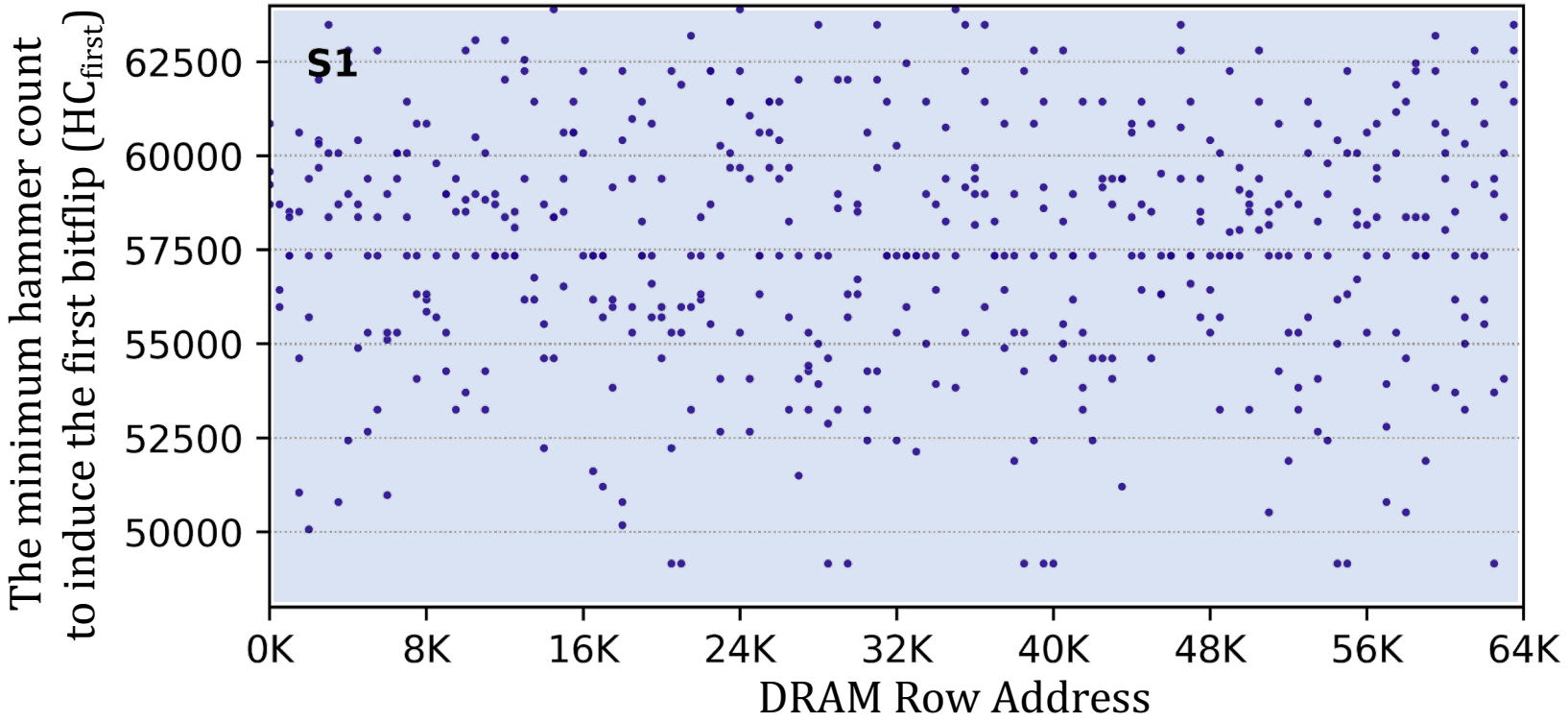
No visible regularity in HC_{first} based on the row's location



The minimum hammer count to induce the first bitflip **irregularly varies** with respect to row's location in DRAM bank

Regularity in Spatial Variation of Read Disturbance across DRAM Rows

No visible regularity in HC_{first} based on the row's location



The minimum hammer count to induce the first bitflip **irregularly varies** with respect to row's location in DRAM bank

Predictability of Read Disturbance Vulnerability

Predictability of a DRAM row's read disturbance vulnerability based on the row's spatial features

- bank address bits
- subarray address bits
- row address bits
- row's distance to local row buffer

Methodology

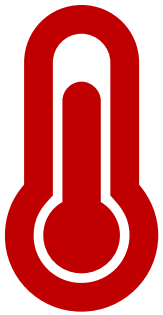
- Cluster DRAM rows into 15 bins based on each row's minimum hammer count to induce the first bitflip (HC_{first})
- Predict whether a row is in a cluster, based on each spatial feature
- Measure the F1 score for these predictions

Key Result: Only a few spatial features have F1 scores > 0.7 only for Mfr. S

A row's **spatial features** are weak predictors for the row's **read disturbance vulnerability**

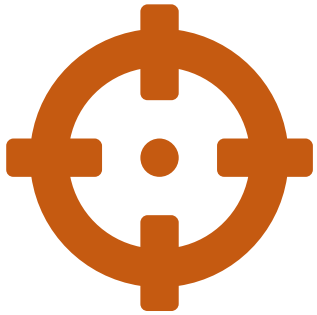
My Dissertation Works

- A deeper look into DRAM read disturbance

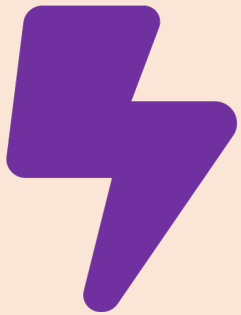


Temperature

Memory Access Patterns



In-Chip Variations



Voltage

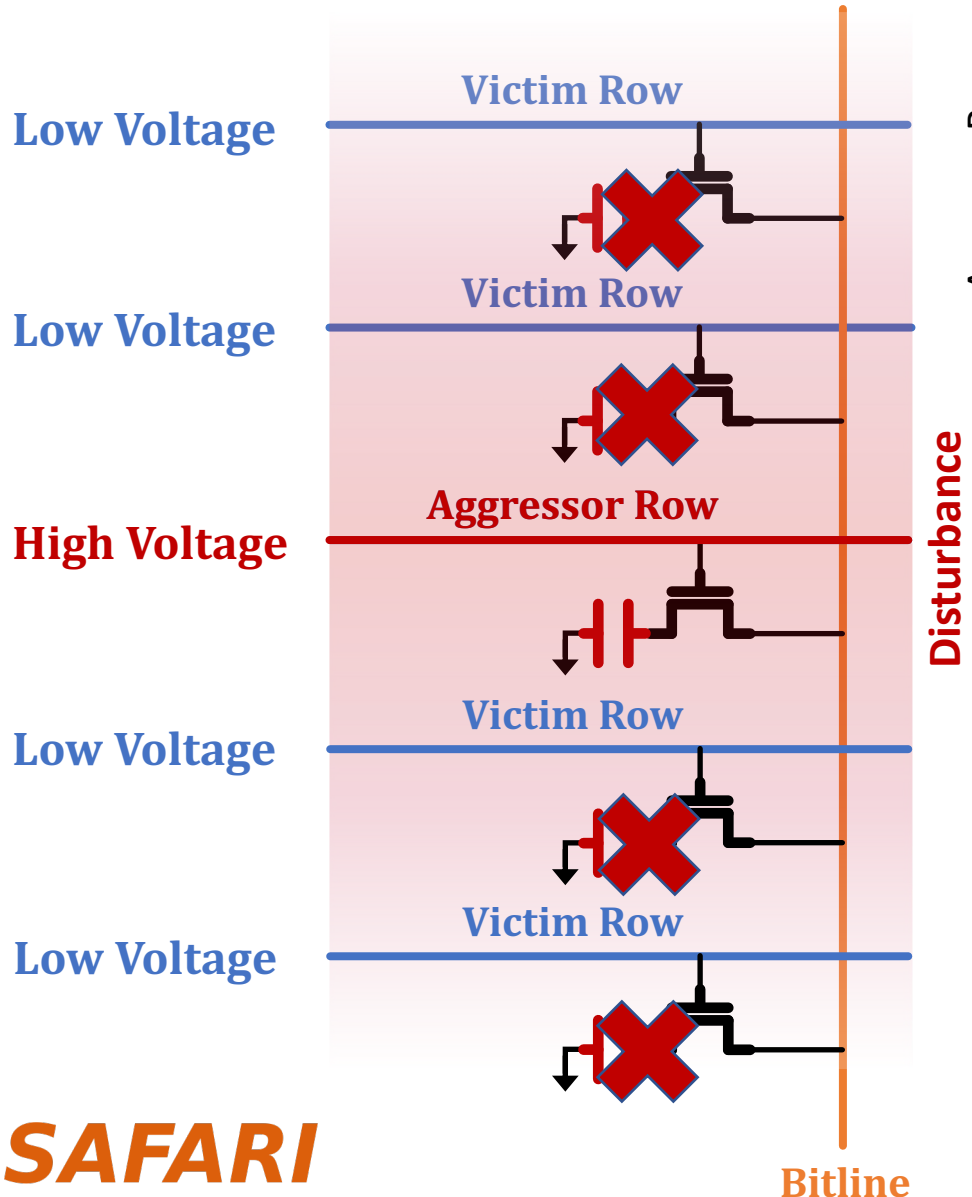
- Solutions to DRAM read disturbance

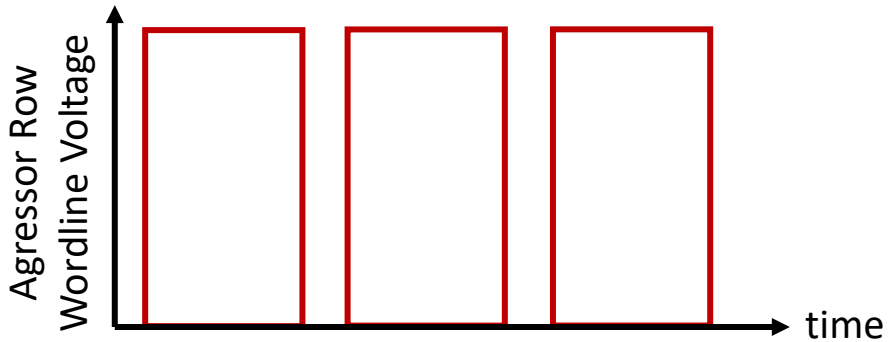
Throttling Unsafe Accesses

Parallelizing Preventive Measures

Leveraging Heterogeneity

A Closer Look into RowHammer



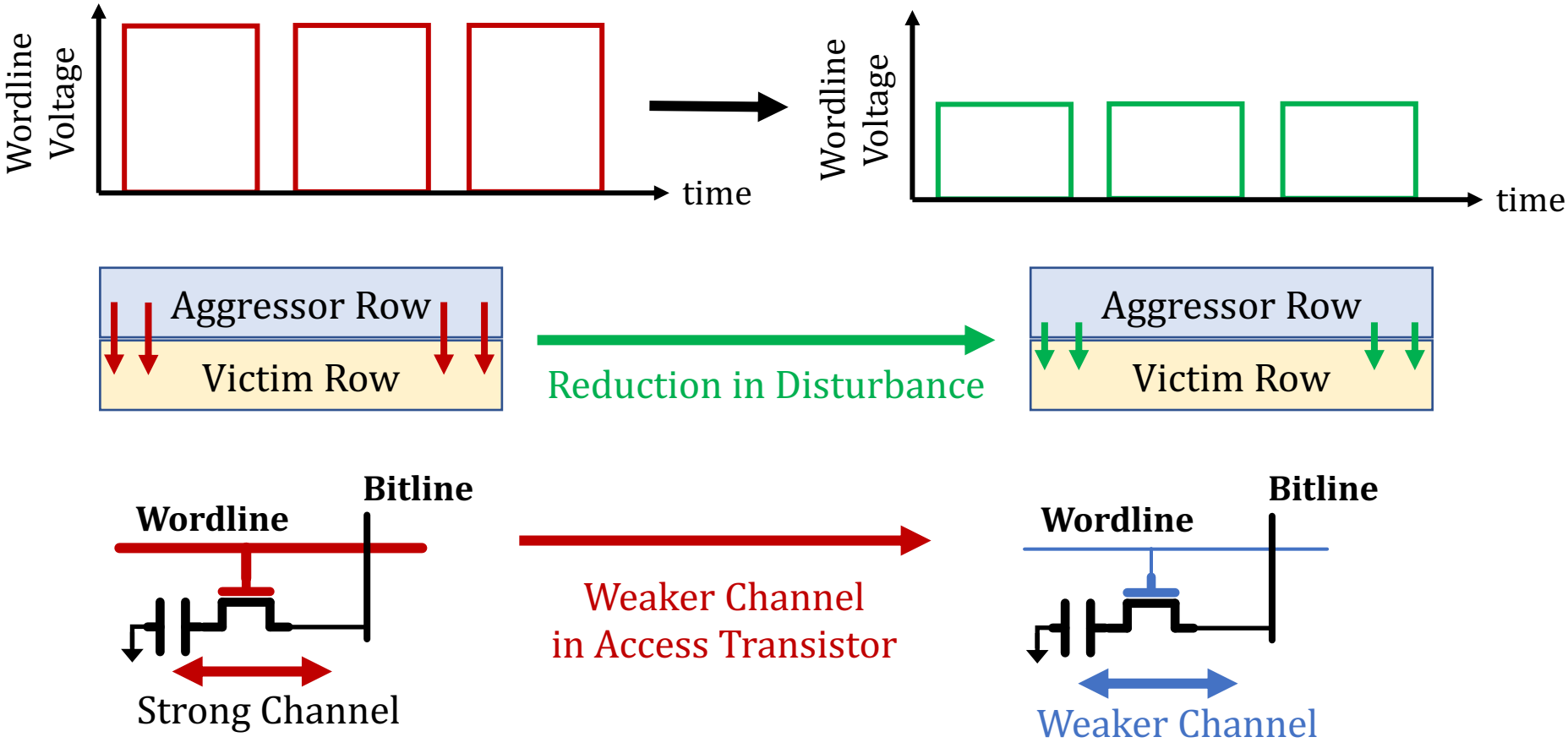


Repeatedly toggling wordline voltage is the *key* to inducing **RowHammer bitflips**

SAFARI

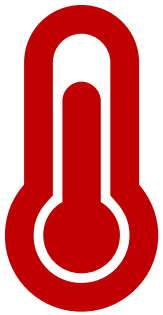
Our Key Finding

Reducing **wordline voltage**
reduces RowHammer vulnerability
without significantly affecting **reliable DRAM operation**



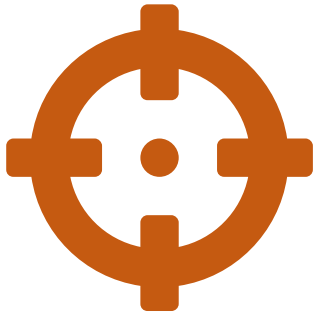
My Dissertation Works

- A deeper look into DRAM read disturbance



Temperature

Memory Access Patterns



In-Chip Variations

Voltage

- Solutions to DRAM read disturbance

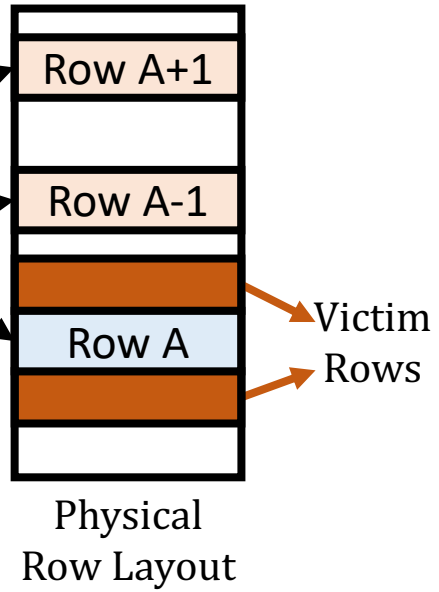
Throttling Unsafe Accesses

Parallelizing Preventive Measures

Leveraging Heterogeneity

Compatibility with Commodity DRAM Chips

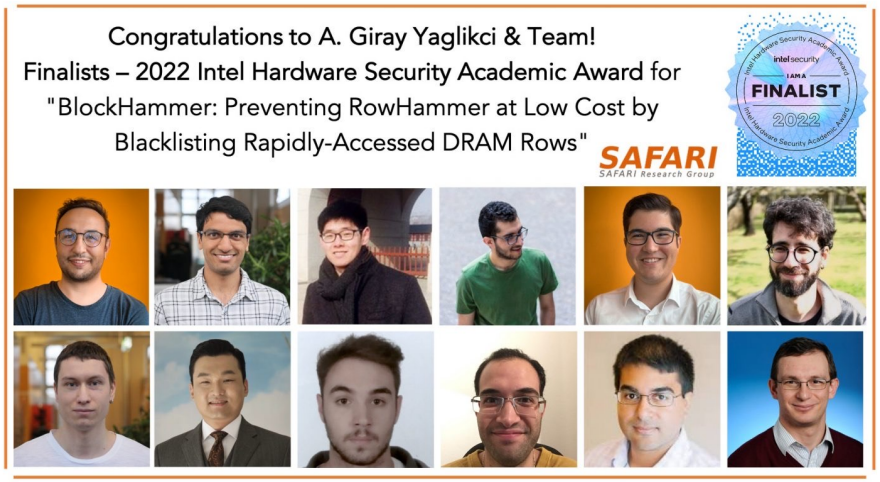
- A **RowHammer attack** hammers Row A
- Existing mechanisms **detect the attack**
- Refresh rows **A+1** and **A-1**
- Bit flips **still may occur** due to **unknown DRAM-internal row mapping**



Existing read disturbance mitigation mechanisms need to know proprietary DRAM-internal row address mapping

BlockHammer – HPCA 2021

- A. Giray Yaglikci, Minesh Patel, Jeremie S. Kim, Roknoddin Azizi, Ataberk Olgun, Lois Orosa, Hasan Hassan, Jisung Park, Konstantinos Kanellopoulos, Taha Shahroodi, Saugata Ghose, and Onur Mutlu,
  "BlockHammer: Preventing RowHammer at Low Cost by Blacklisting Rapidly-Accessed DRAM Rows"
  Proceedings of the 27th International Symposium on High-Performance Computer Architecture (***HPCA***),
  Virtual, February-March 2021.
  [Slides (pptx) (pdf)]
  [Short Talk Slides (pptx) (pdf)]
  [Intel Hardware Security Academic Awards Short Talk Slides (pptx) (pdf)]
  [Talk Video (22 minutes)]
  [Short Talk Video (7 minutes)]
  [Intel Hardware Security Academic Awards Short Talk Video (2 minutes)]
  [BlockHammer Source Code]
  Intel Hardware Security Academic Award Finalist (one of 4 finalists out of 34 nominations)

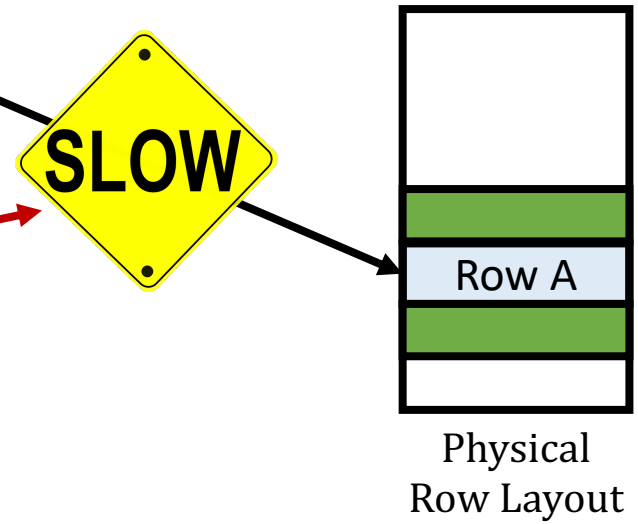


BlockHammer: Preventing RowHammer at Low Cost by Blacklisting Rapidly-Accessed DRAM Rows

A. Giray Yağlıkçı[1] Minesh Patel[1] Jeremie S. Kim[1] Roknoddin Azizi[1] Ataberk Olgun[1] Lois Orosa[1]
Hasan Hassan[1] Jisung Park[1] Konstantinos Kanellopoulos[1] Taha Shahroodi[1] Saugata Ghose[2] Onur Mutlu[1]
[1]*ETH Zürich* [2]*University of Illinois at Urbana–Champaign*

BlockHammer: Throttling Unsafe Accesses

- A RowHammer attack hammers Row A
- BlockHammer detects and **selectively throttles accesses** from within **the memory controller**
- Bit flips **do not** occur



Physical Row Layout

- BlockHammer can *optionally* **inform the system software** about the attack

BlockHammer is compatible with commodity DRAM chips
No need for proprietary info of or modifications to DRAM chips

Spatial Variation-Aware Read Disturbance Defenses: Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

Aging Slides

Full Paper
arXiv [cs.CR] 2402.18652

Abdullah Giray Yağlıkçı

Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

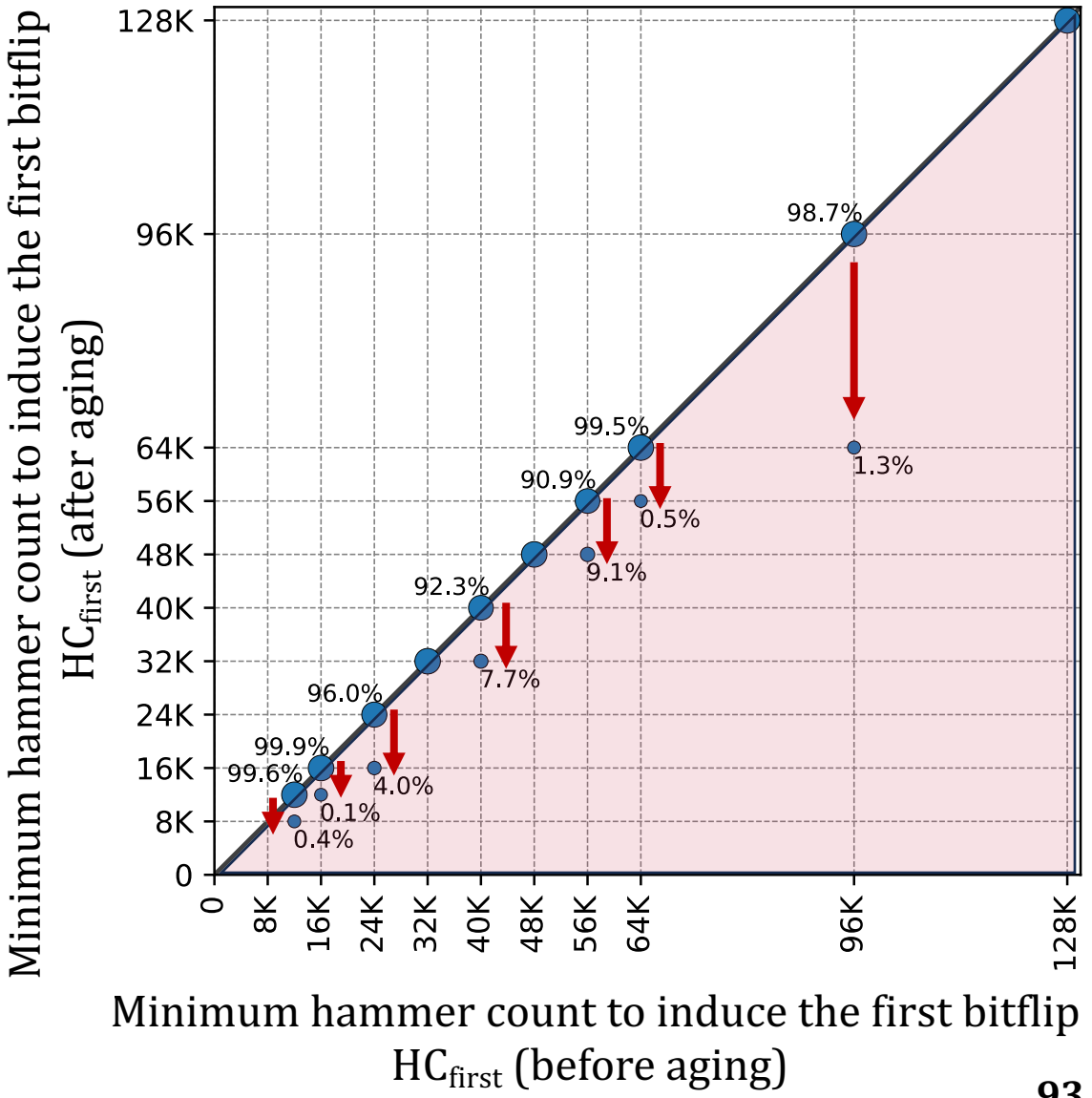
Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI ETH zürich

More in the Paper (2/2): Aging Study

Preliminary data on aging via 68-day of continuous hammering

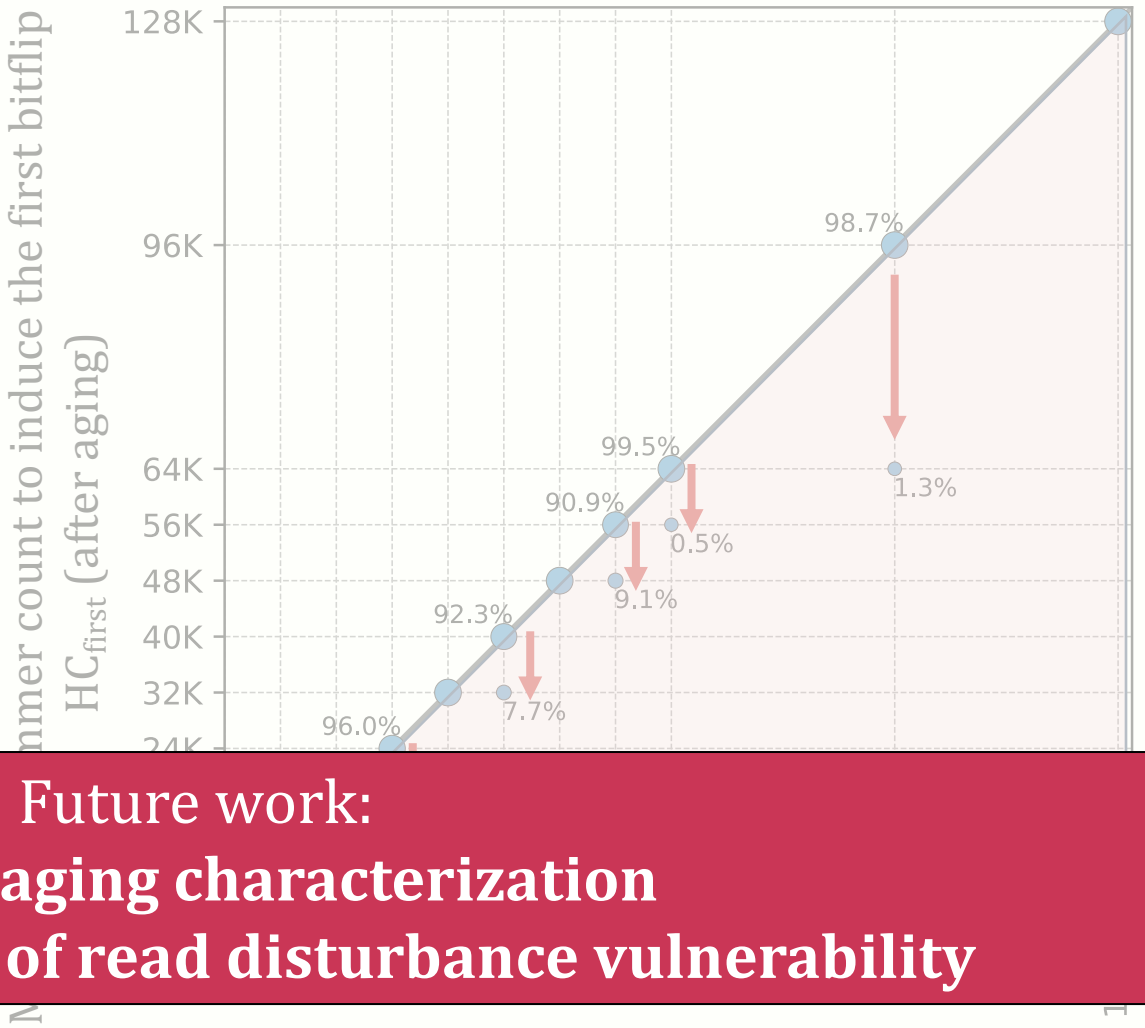
Aging can lead to read disturbance bitflips at **smaller** hammer counts



More in the Paper (2/2): Aging Study

Preliminary data on aging via 68-day of continuous hammering

Aging can lead to read disturbance bitflips at **smaller** hammer counts



Future work:
rigorous aging characterization
and **online profiling of read disturbance vulnerability**

Spatial Variation-Aware Read Disturbance Defenses: Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

Hydra's Performance

Full Paper
arXiv [cs.CR] 2402.18652

Abdullah Giray Yağlıkçı

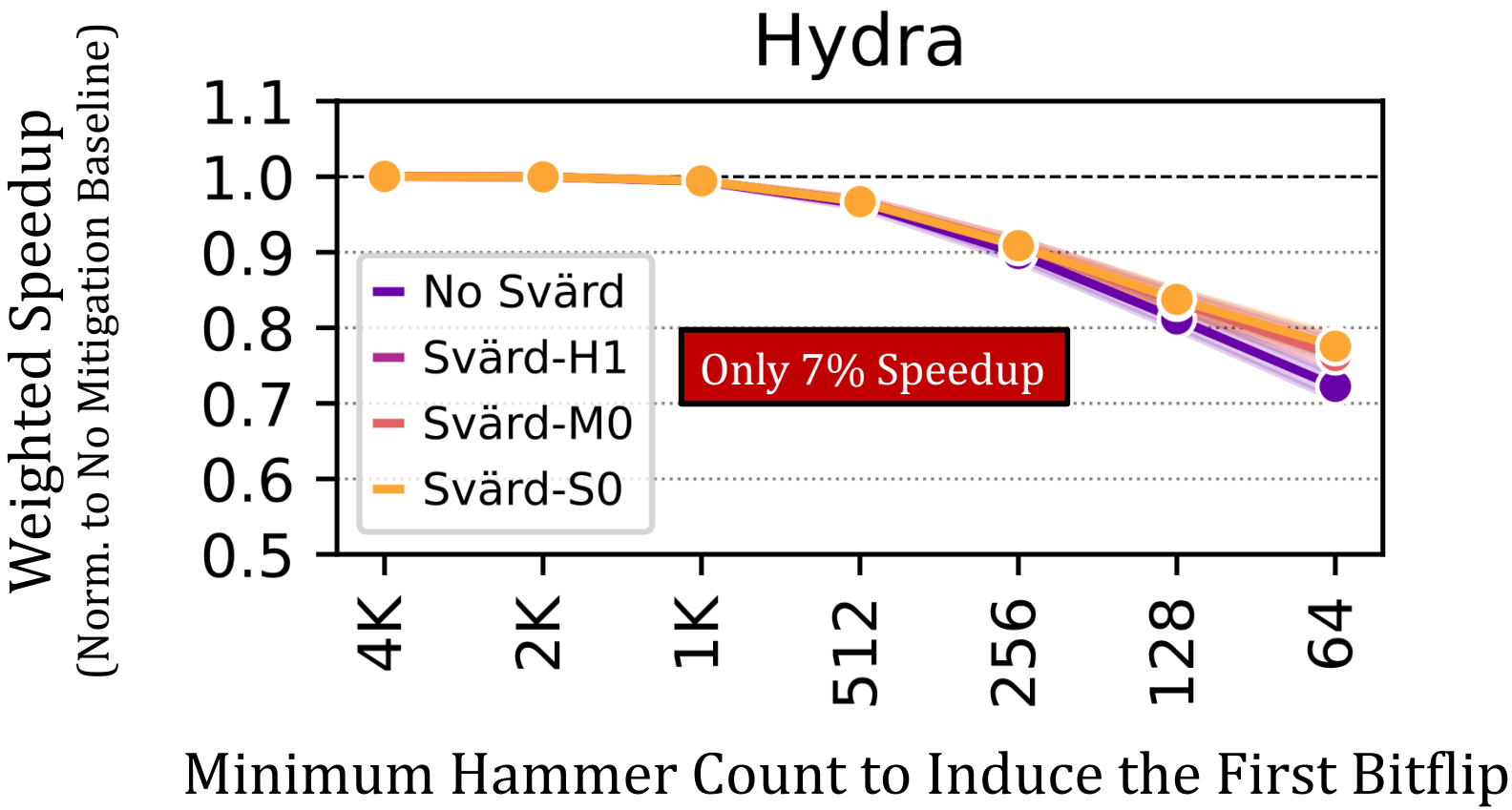
Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI

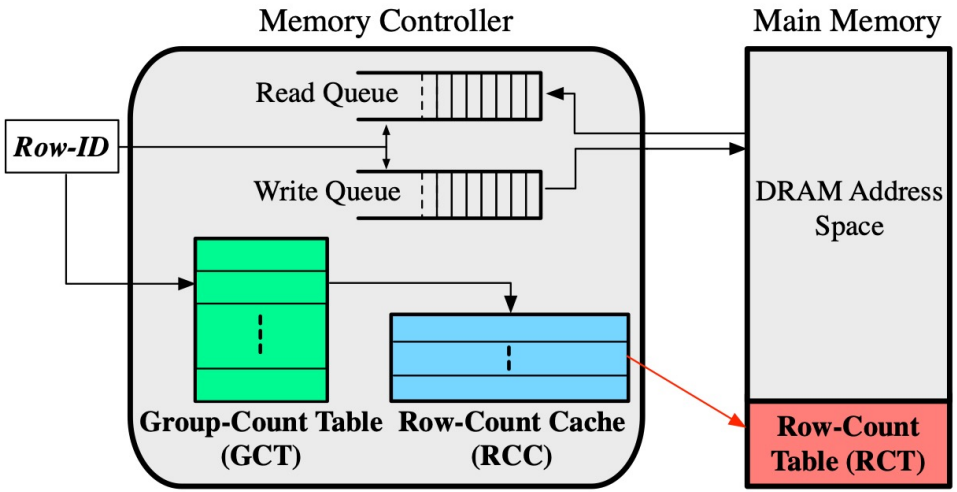
ETH zürich

An Outlier Solution: Hydra



Hydra Mitigation Mechanism

- Hydra maintains **a row activation counter for each DRAM row**
- **Stores** these activation counters **in the DRAM array**
- **Caches** the counters of hot rows **in the memory controller**
- At low HC_{first} configurations, **many rows are hot**
- **Fetching / evicting counters** dominate the performance overhead
- Svärd needs **further customizations** for Hydra



Qureshi, et al. **"Hydra: Enabling Low-Overhead Mitigation of Row-Hammer at Ultra-Low Thresholds via Hybrid Tracking," in ISCA, 2022.**

Spatial Variation-Aware Read Disturbance Defenses: Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

HC_{first} across Rows

Full Paper
arXiv [cs.CR] 2402.18652

Abdullah Giray Yağlıkçı

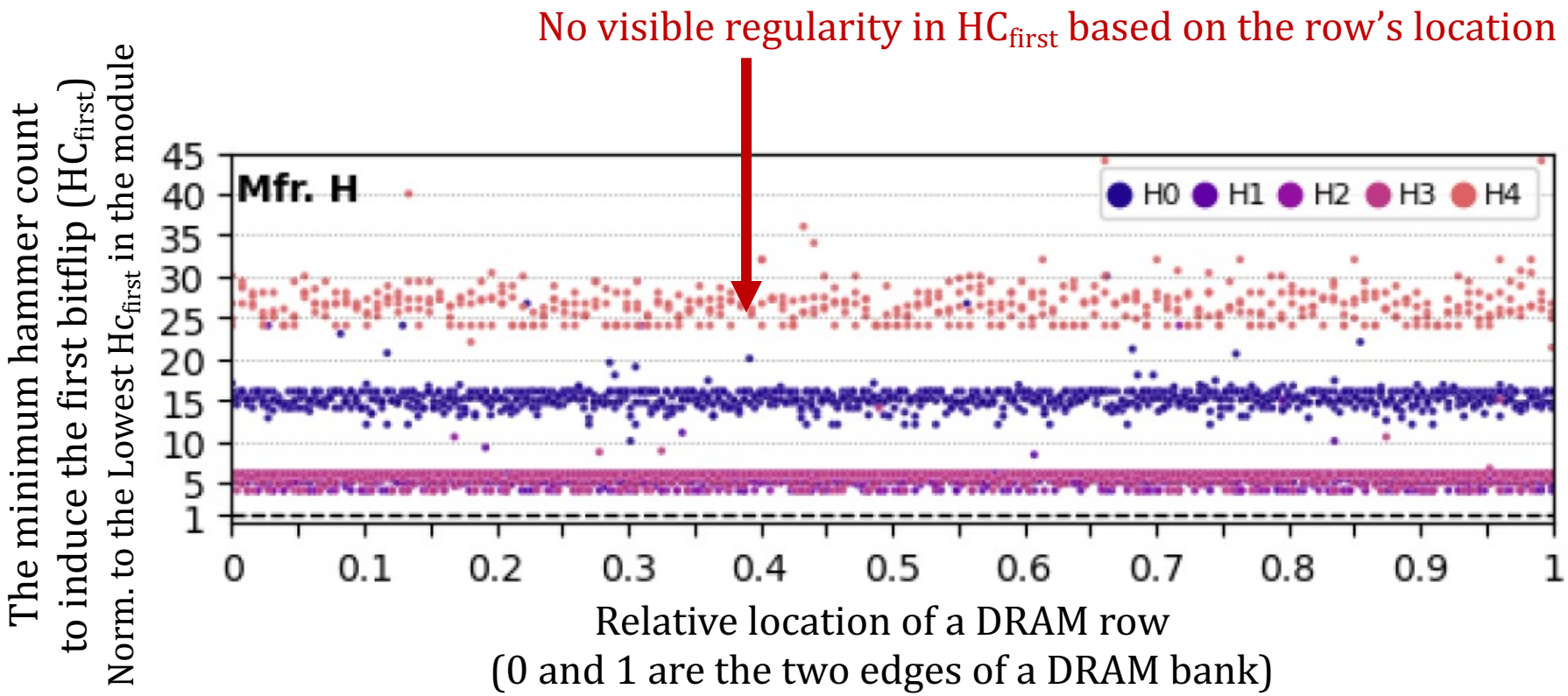
Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI

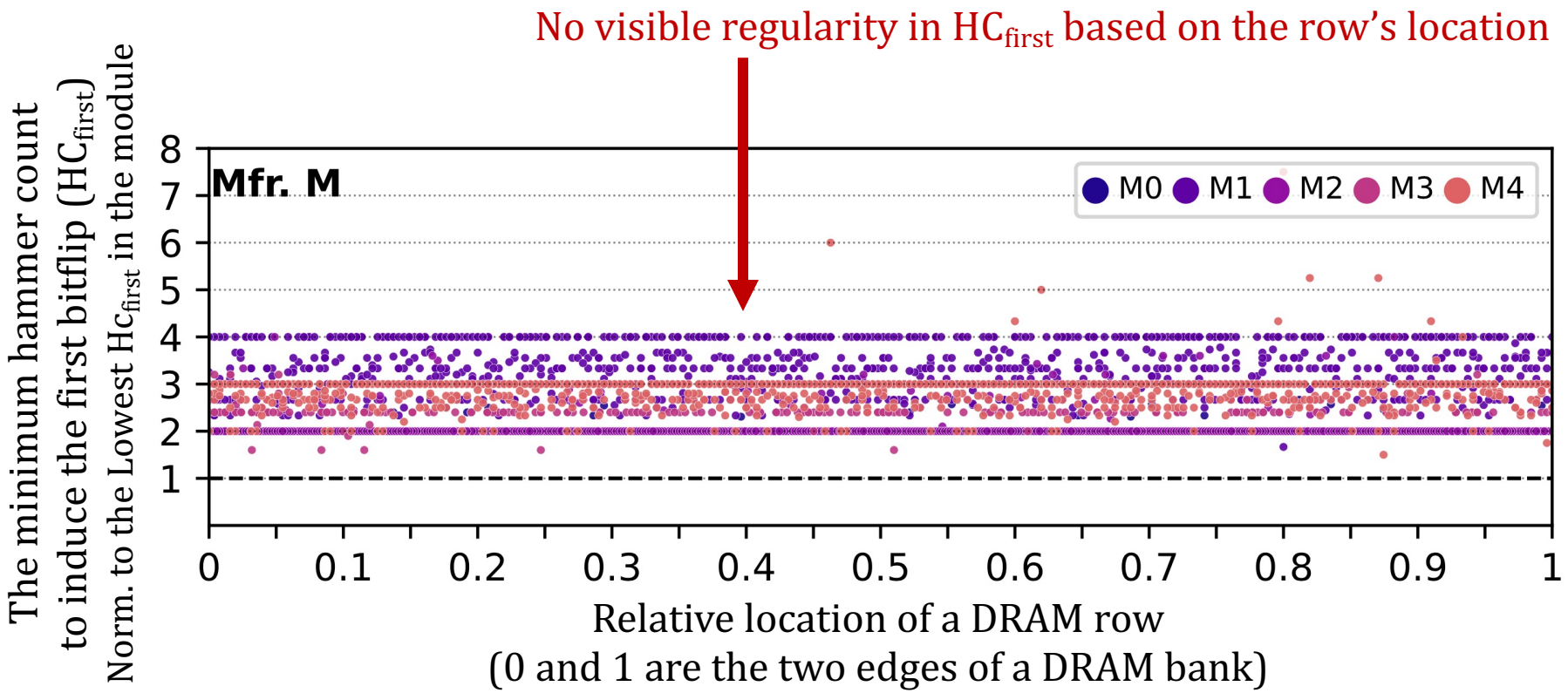
ETH zürich

The Minimum Hammer Count to Induce the First Bitflip across DRAM Rows



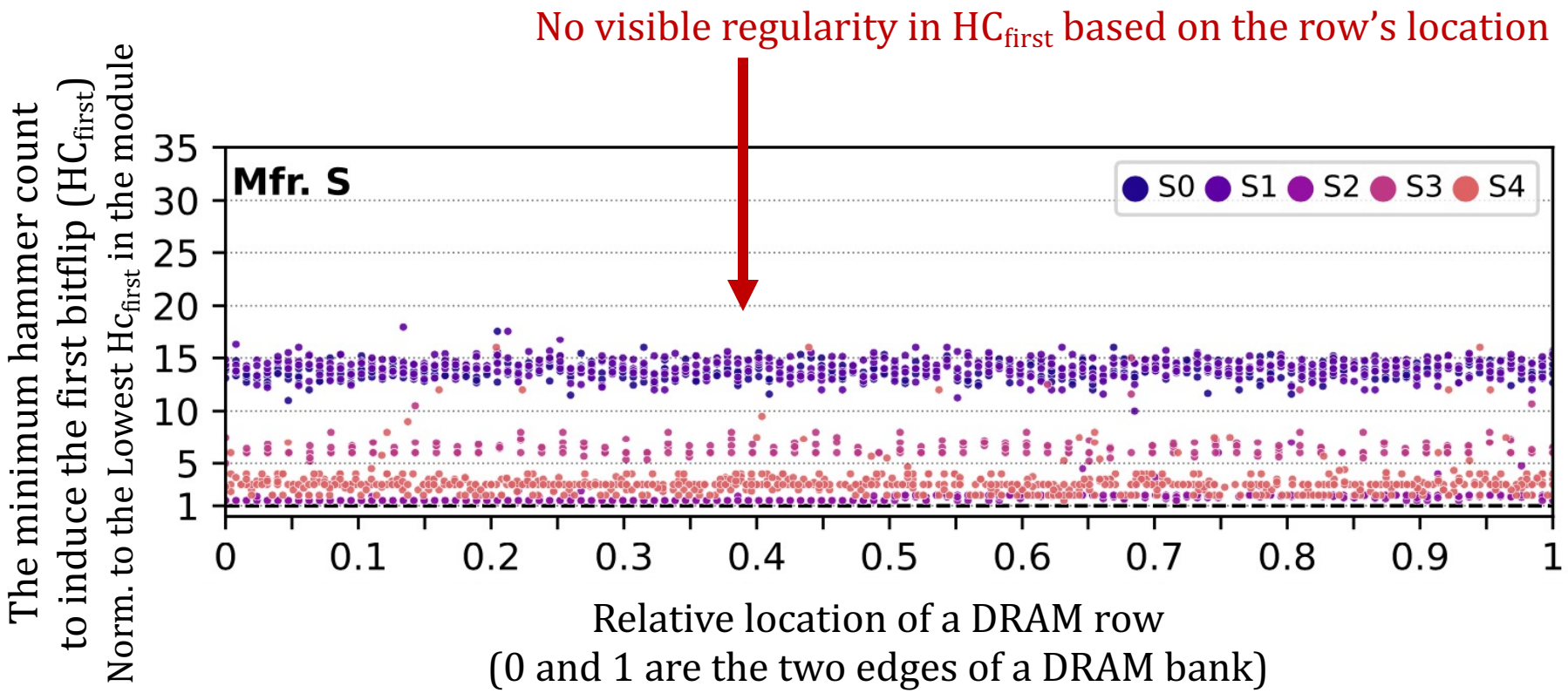
The minimum hammer count to induce the first bitflip **irregularly varies** with respect to row's location in DRAM bank

The Minimum Hammer Count to Induce the First Bitflip across DRAM Rows



The minimum hammer count to induce the first bitflip **irregularly varies** with respect to row's location in DRAM bank

The Minimum Hammer Count to Induce the First Bitflip across DRAM Rows



The minimum hammer count to induce the first bitflip **irregularly varies** with respect to row's location in DRAM bank

Spatial Variation-Aware Read Disturbance Defenses:
Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

BER across Rows

Full Paper
arXiv [cs.CR] 2402.18652

Abdullah Giray Yağlıkçı

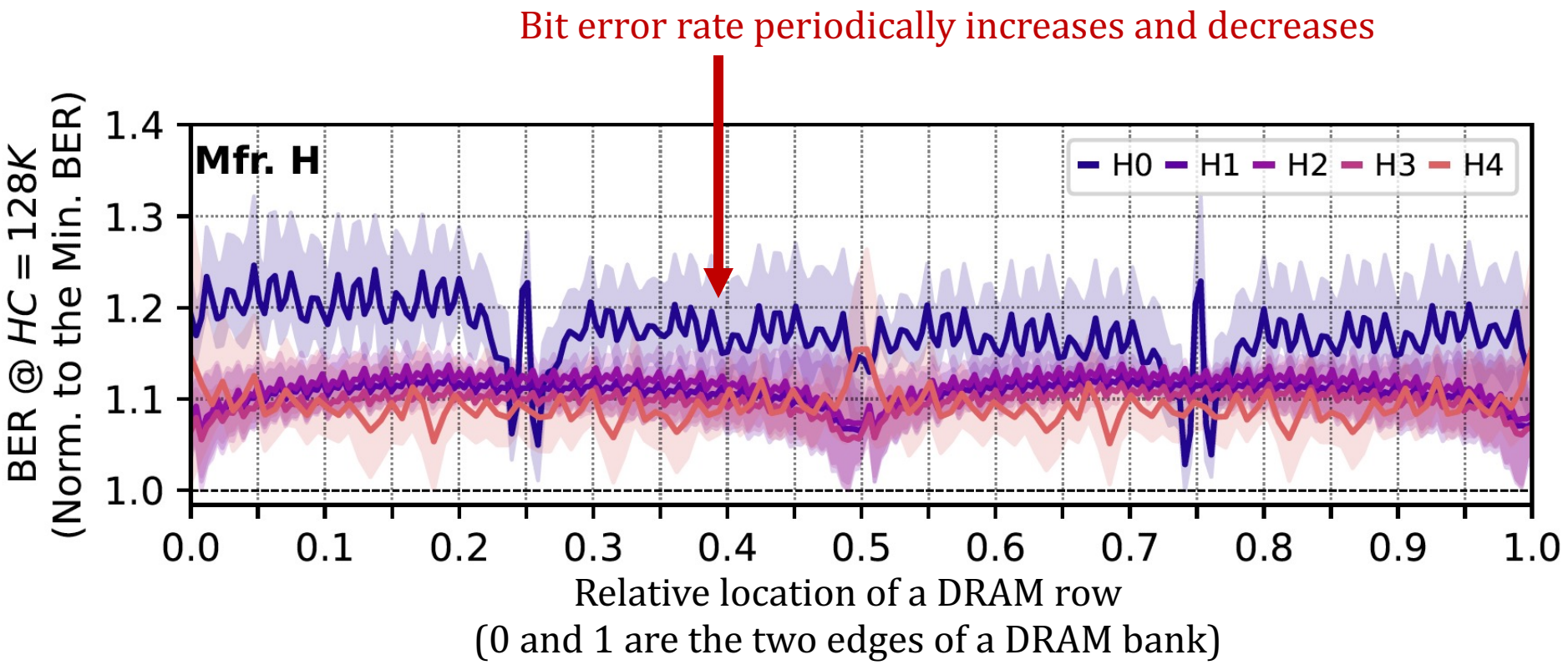
Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI

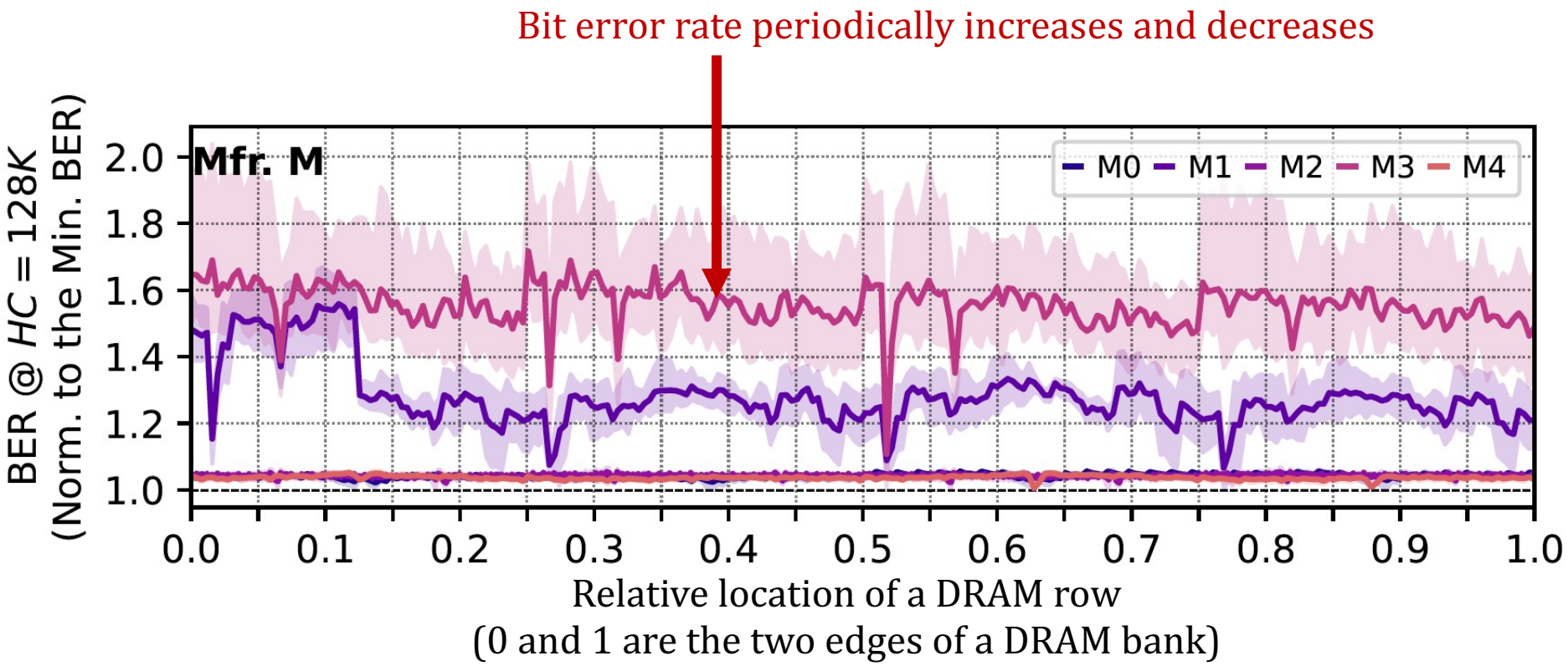
ETH zürich

The Read Disturbance Bit Error Rate across DRAM Rows



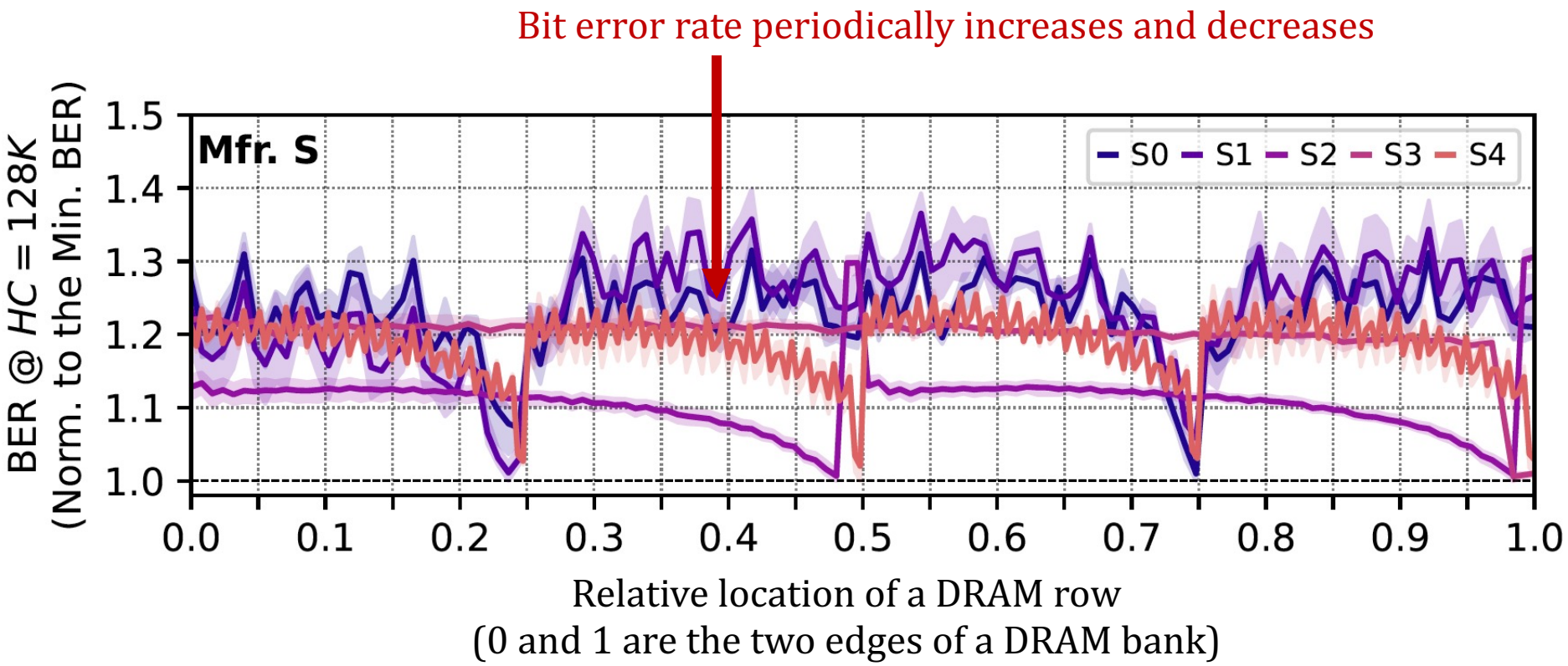
The variation in read disturbance bit error rate **exhibits a stronger regularity compared to the variation in varies** with respect to row's location in DRAM bank

The Read Disturbance Bit Error Rate across DRAM Rows



The variation in read disturbance bit error rate **exhibits a stronger regularity compared to the variation in varies** with respect to row's location in DRAM bank

The Read Disturbance Bit Error Rate across DRAM Rows



The variation in read disturbance bit error rate **exhibits a stronger regularity compared to the variation in varies** with respect to row's location in DRAM bank

Spatial Variation-Aware Read Disturbance Defenses: Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

Background

Full Paper
arXiv [cs.CR] 2402.18652

Abdullah Giray Yağlıkçı

Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

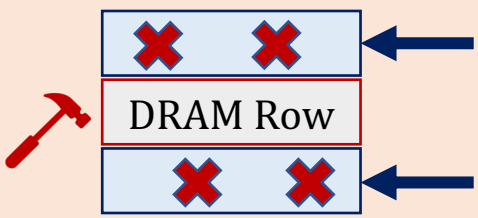
Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI ETH zürich

Two Main Types of DRAM Refresh

1 **Periodic Refresh:** Periodically **restores** the charge DRAM cells leak **over time**

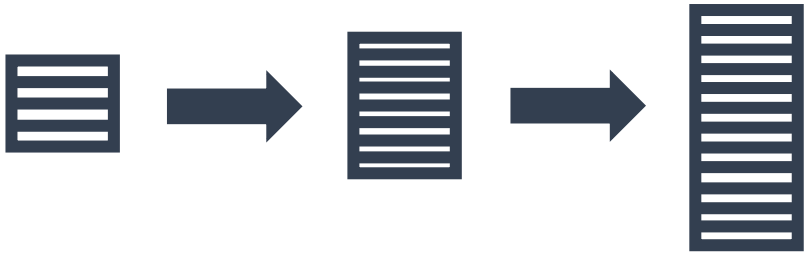
2 **RowHammer:** **Repeatedly accessing** a DRAM row can cause **bit flips** in other **physically nearby rows**



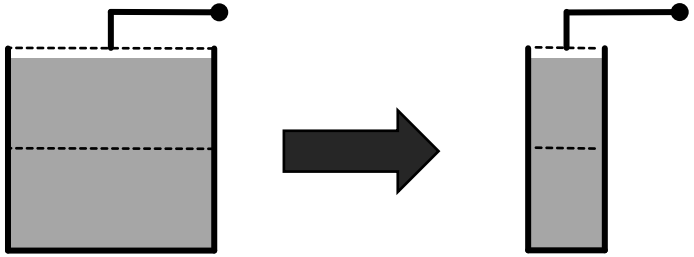
Preventive Refresh: Mitigates RowHammer by **refreshing physically nearby rows** of a repeatedly accessed row

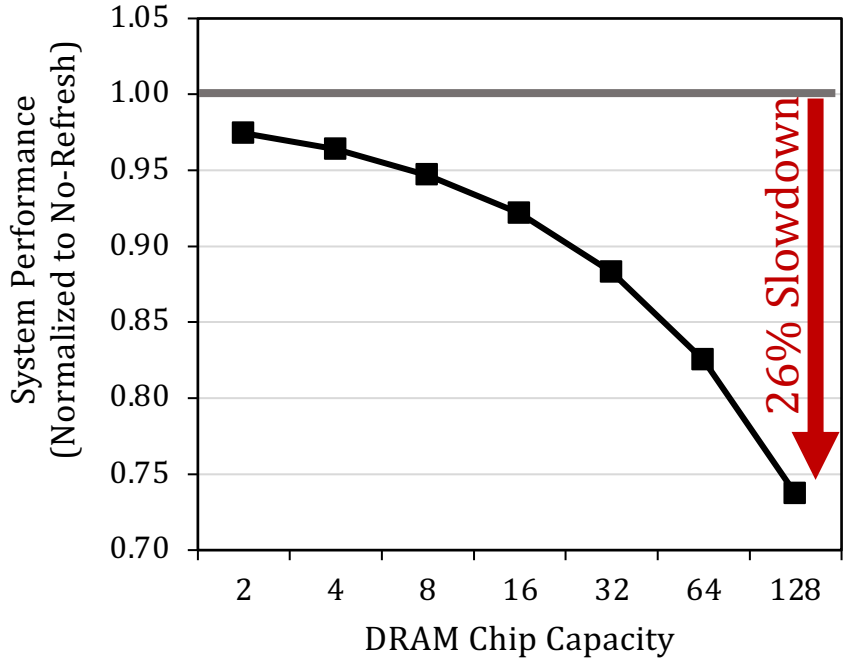
Periodic Refresh with Increasing DRAM Chip Density

A **larger capacity** chip has **more rows to be refreshed**



A **smaller** cell stores **less charge**





More periodic refresh operations incur **larger performance overhead** as DRAM **chip density increases**

Spatial Variation-Aware Read Disturbance Defenses: Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

RowPress

Full Paper
arXiv [cs.CR] 2402.18652

Abdullah Giray Yağlıkçı

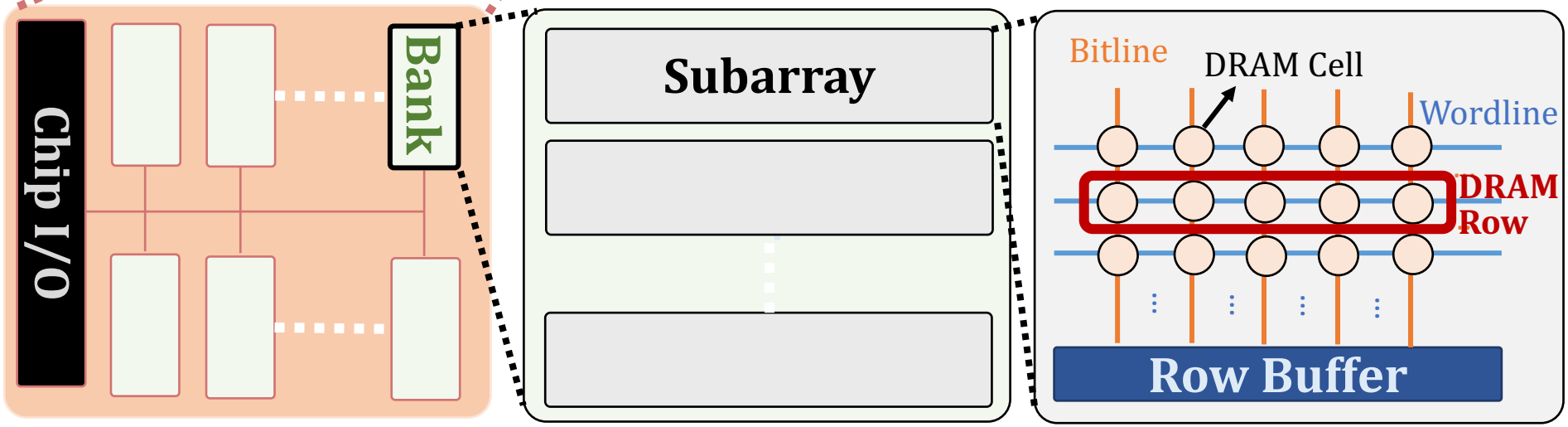
Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI

ETH zürich

DRAM Organization

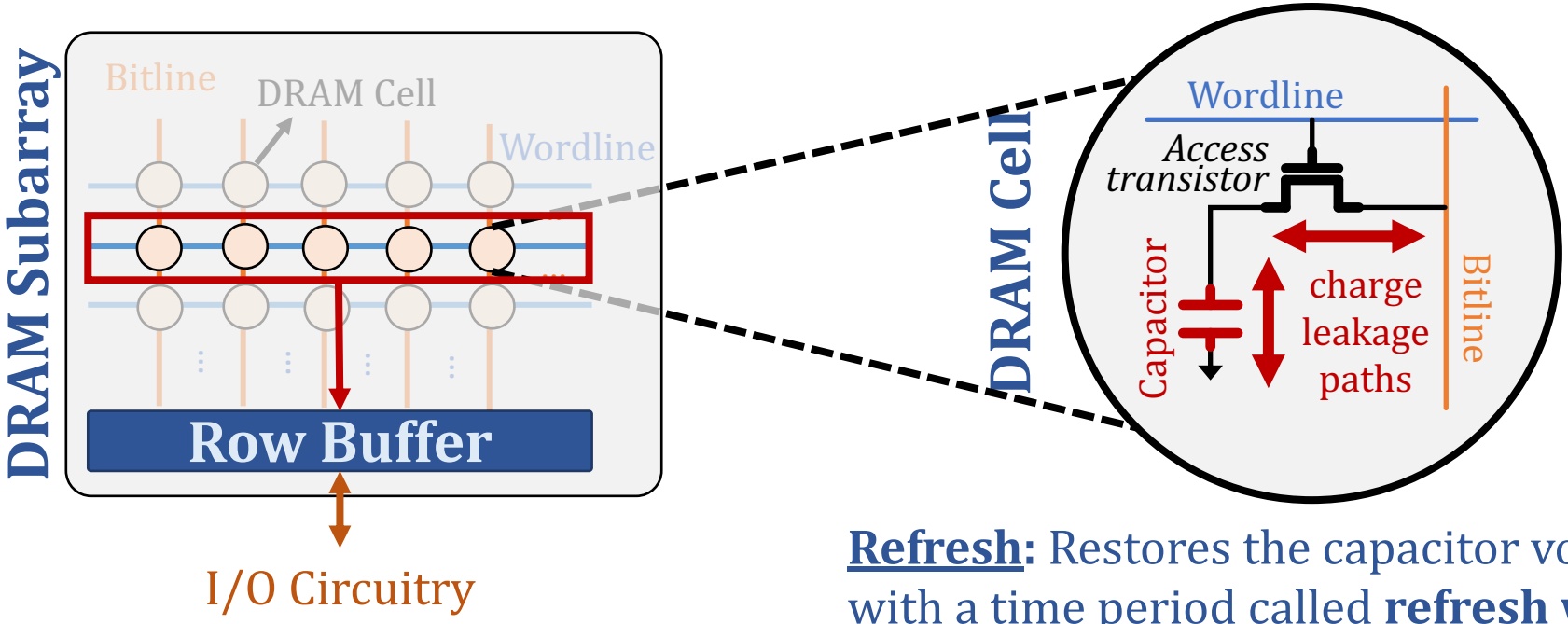


DRAM Chip

DRAM Bank

DRAM Subarray

DRAM Operation

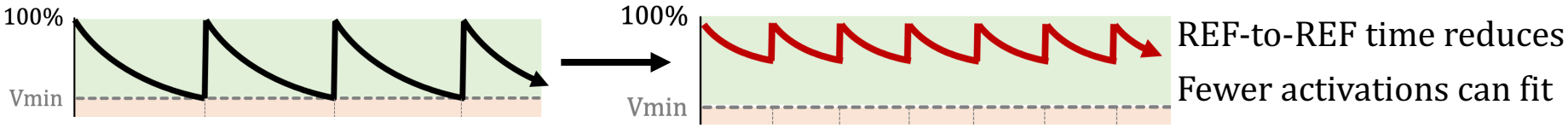


Refresh: Restores the capacitor voltage with a time period called **refresh window**

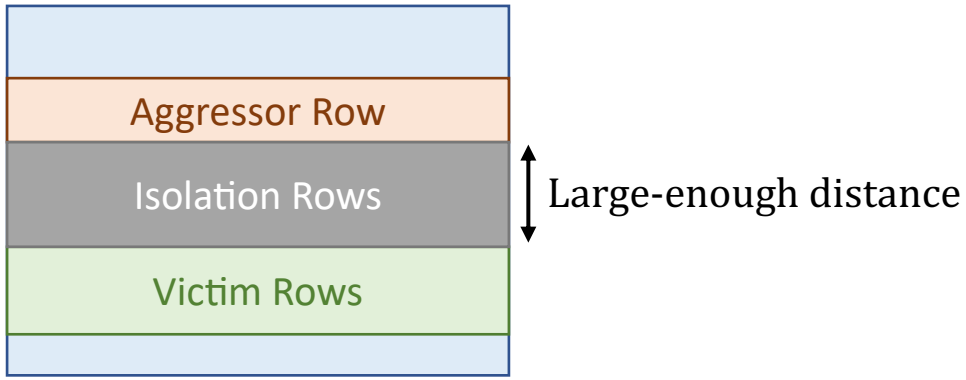
1. **Row Activation:** Fetch the row's content into the row buffer
2. **Column Access:** Read/Write a column in the row buffer
3. **Precharge:** Disconnect the row from the row buffer

RowHammer Mitigation Approaches

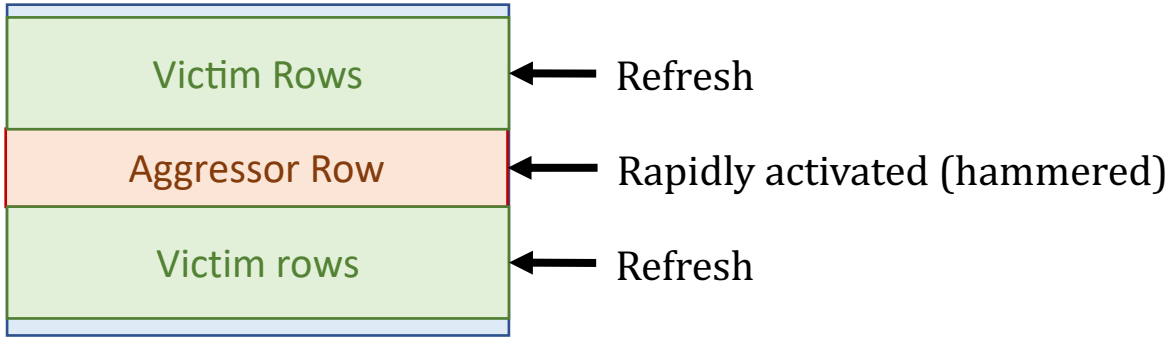
- Increased refresh rate



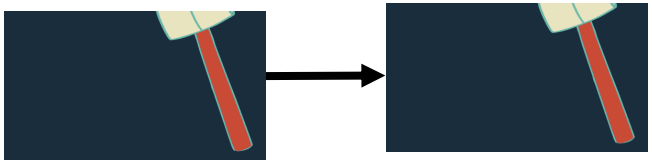
- Physical isolation



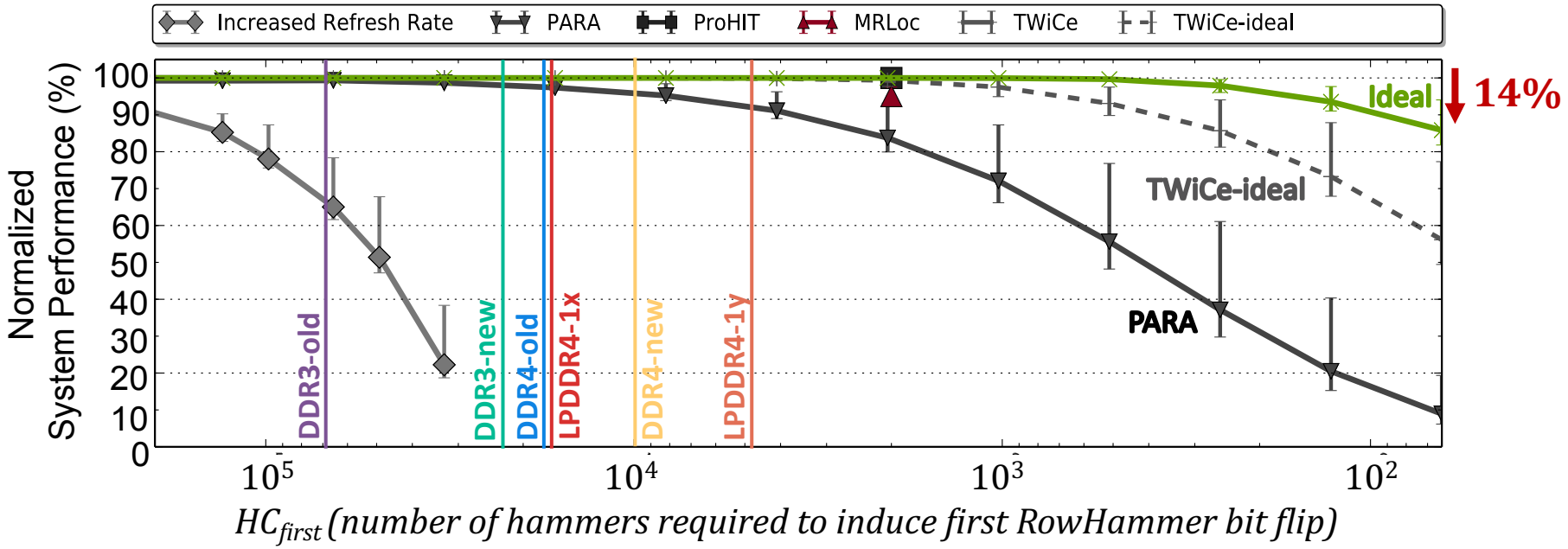
- Reactive refresh



- Proactive throttling



RowHammer Mitigation across Generations



J. S. Kim, M. Patel, A. G. Yaglikci, H. Hassan, R. Azizi, L. Orosa, and O. Mutlu, "Revisiting RowHammer: An Experimental Analysis of Modern Devices and Mitigation Techniques," in *ISCA*, 2020.

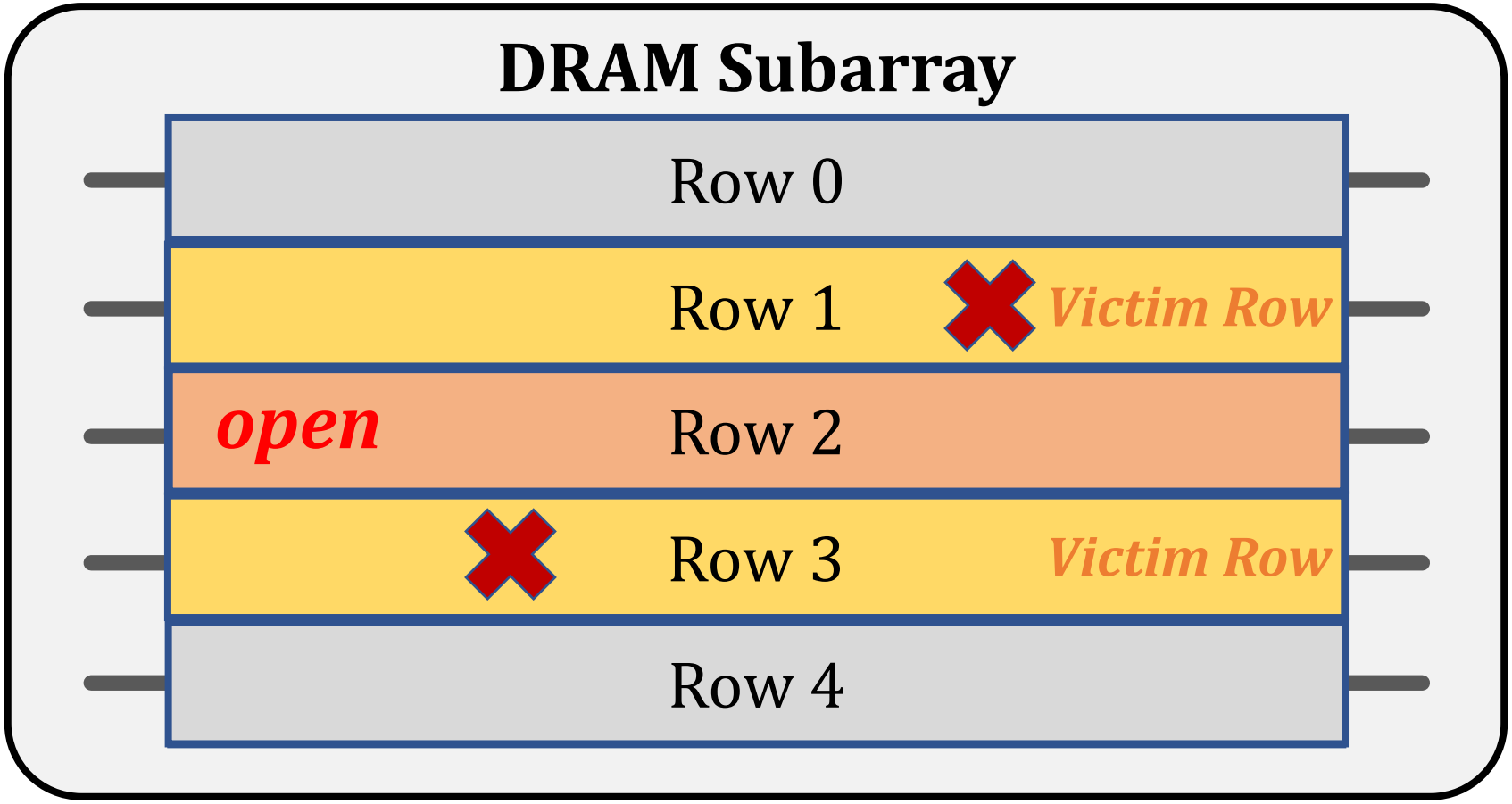
RowPress [ISCA 2023]

- Haocong Luo, Ataberk Olgun, Giray Yaglikci, Yahya Can Tugrul, Steve Rhyner, M. Banu Cavlak, Joel Lindegger, Mohammad Sadrosadati, and Onur Mutlu,
  "RowPress: Amplifying Read Disturbance in Modern DRAM Chips"
  Proceedings of the 50th International Symposium on Computer Architecture (***ISCA***), Orlando, FL, USA, June 2023.
  [Slides (pptx) (pdf)]
  [Lightning Talk Slides (pptx) (pdf)]
  [Lightning Talk Video (3 minutes)]
  [RowPress Source Code and Datasets (Officially Artifact Evaluated with All Badges)]
  Officially artifact evaluated as available, reusable and reproducible.
  Best artifact award at ISCA 2023.

RowPress: Amplifying Read-Disturbance in Modern DRAM Chips

Haocong Luo Ataberk Olgun A. Giray Yağlıkçı Yahya Can Tuğrul Steve Rhyner
Meryem Banu Cavlak Joël Lindegger Mohammad Sadrosadati Onur Mutlu
ETH Zürich

RowPress A New DRAM Read Disturbance Phenomenon



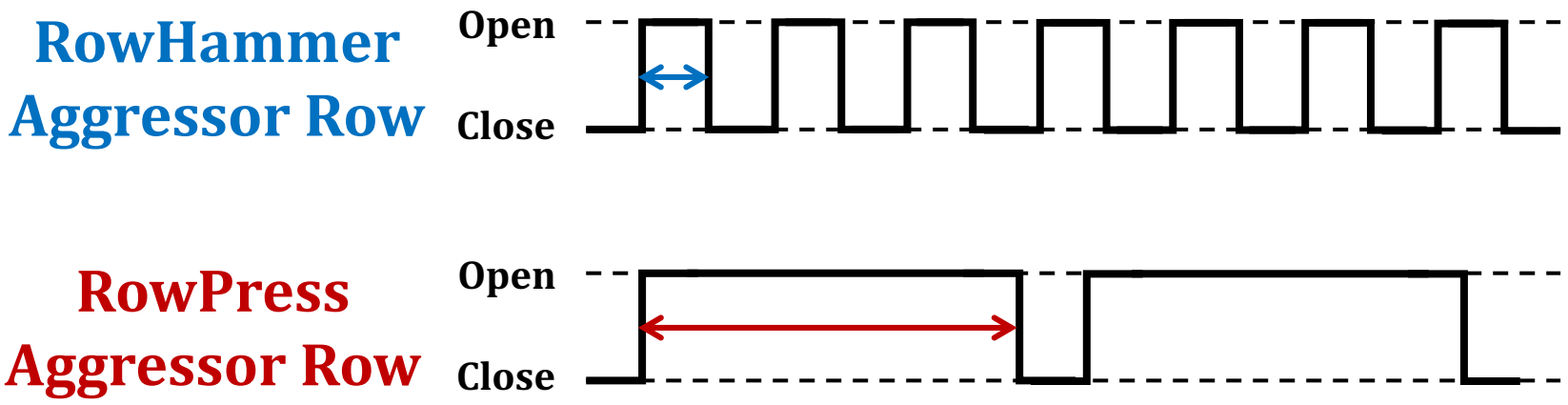
Keeping a DRAM row **open** (activated)

causes **RowPress bitflips** in nearby cells

Two Prime Examples of DRAM Read Disturbance: RowHammer and RowPress

Instead of using a high activation count,

☞ increase the time that the aggressor row stays open



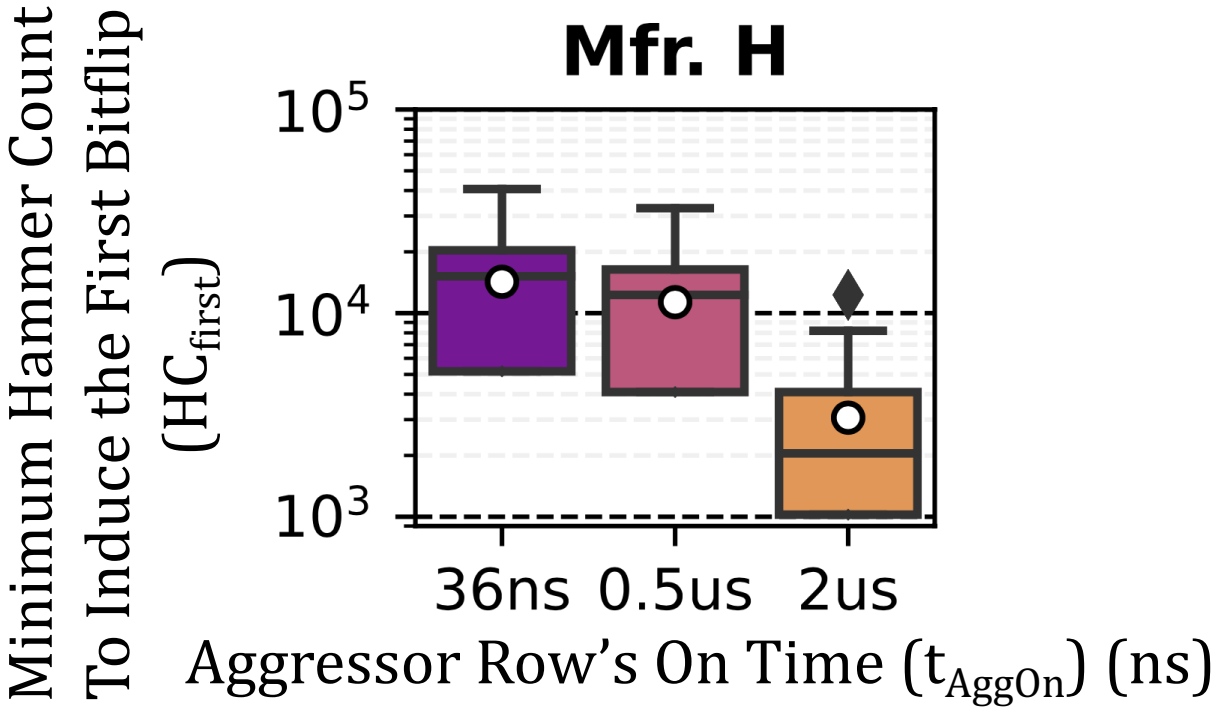
RowPress [ISCA 2023]

- Haocong Luo, Ataberk Olgun, Giray Yaglikci, Yahya Can Tugrul, Steve Rhyner, M. Banu Cavlak, Joel Lindegger, Mohammad Sadrosadati, and Onur Mutlu,
  "RowPress: Amplifying Read Disturbance in Modern DRAM Chips"
  Proceedings of the 50th International Symposium on Computer Architecture (***ISCA***), Orlando, FL, USA, June 2023.
  [Slides (pptx) (pdf)]
  [Lightning Talk Slides (pptx) (pdf)]
  [Lightning Talk Video (3 minutes)]
  [RowPress Source Code and Datasets (Officially Artifact Evaluated with All Badges)]
  Officially artifact evaluated as available, reusable and reproducible.
  Best artifact award at ISCA 2023.

RowPress: Amplifying Read-Disturbance in Modern DRAM Chips

Haocong Luo Ataberk Olgun A. Giray Yağlıkçı Yahya Can Tuğrul Steve Rhyner
Meryem Banu Cavlak Joël Lindegger Mohammad Sadrosadati Onur Mutlu
ETH Zürich

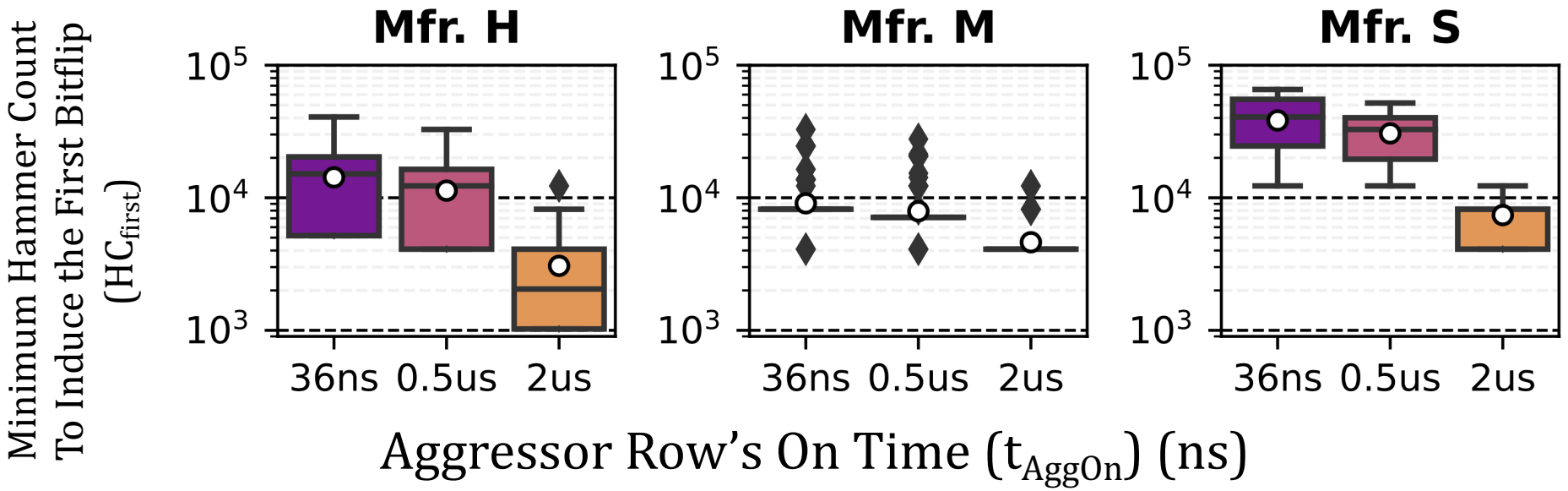
Effect of RowPress on Spatial Variation of Read Disturbance Vulnerability across Rows



RowPress reduces the mean of the distribution with increased t_{AggOn}

There is still significant variation in the HC_{first} across rows

Effect of RowPress on the HCfirst Distribution



Bitflips occur at **lower hammer counts** with RowPress
and these hammer counts still **significantly vary** across DRAM rows

Spatial Variation-Aware Read Disturbance Defenses:
Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

Predictability

Full Paper
arXiv [cs.CR] 2402.18652

Abdullah Giray Yağlıkçı

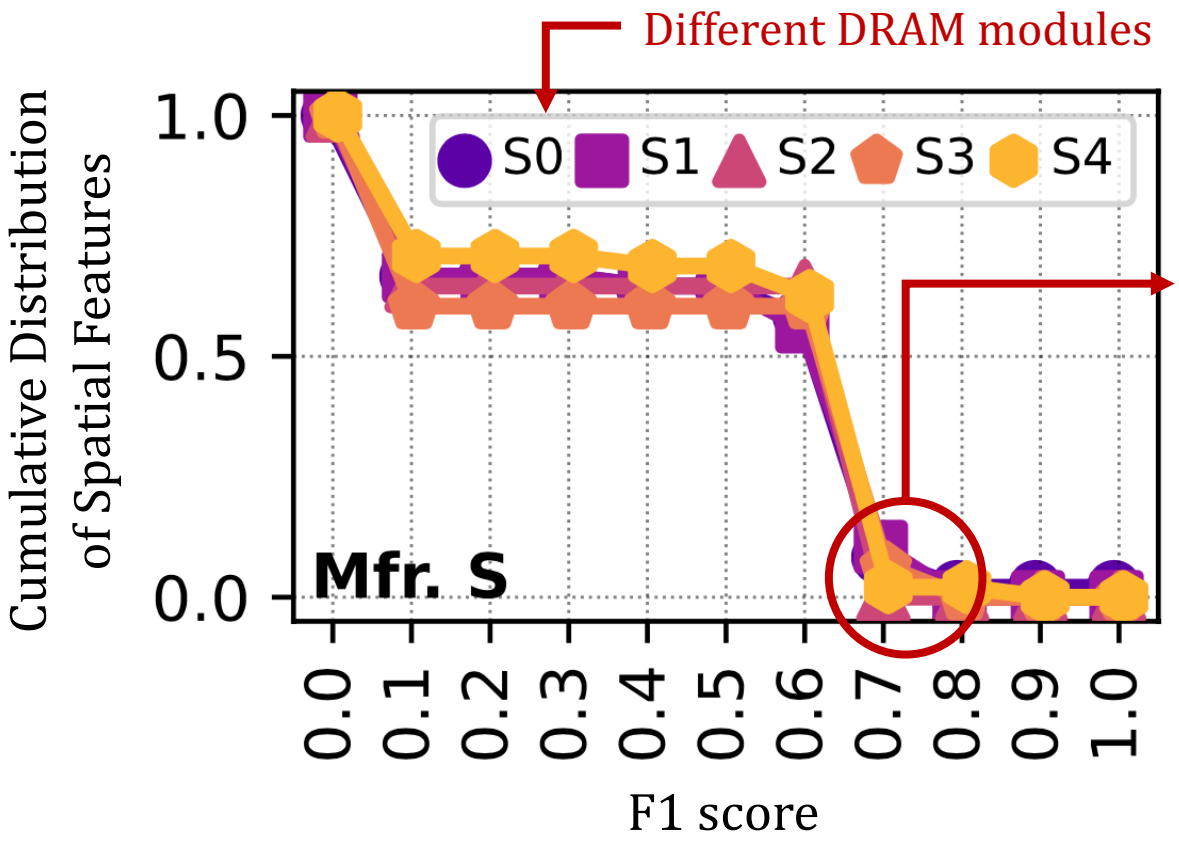
Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI

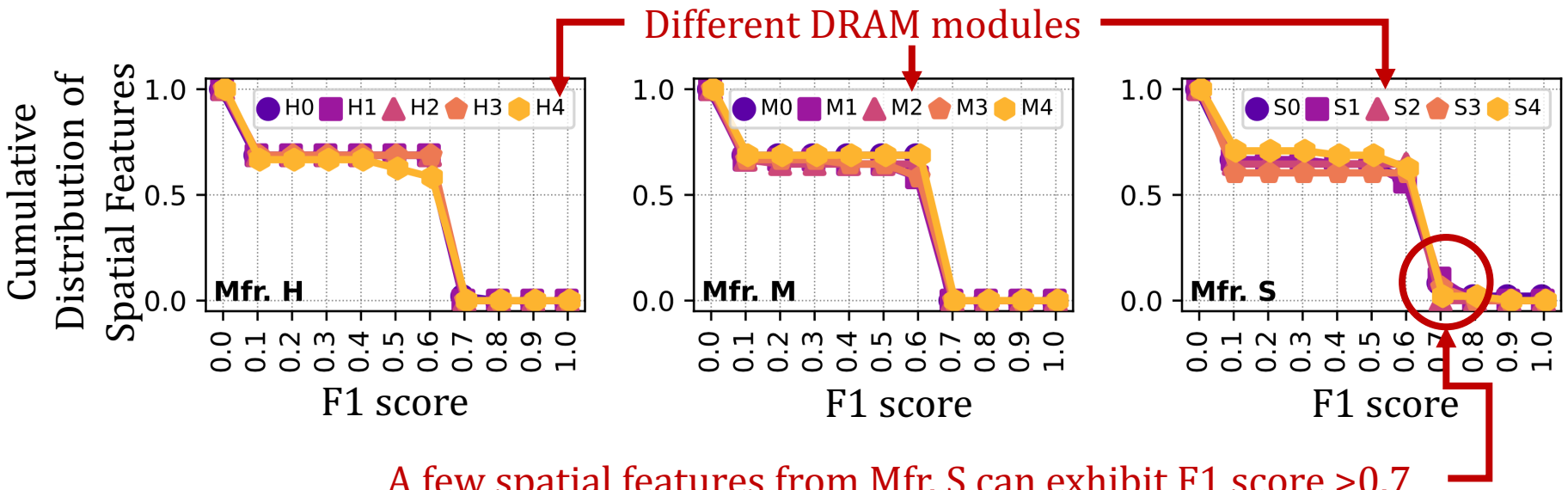
ETH zürich

Cumulative Distribution of Spatial Features based on F1 Score



Weak prediction observed between a row's **spatial features** & read disturbance **vulnerability**

Cumulative Distribution of Spatial Features based on F1 Score



Weak prediction observed between a row's **spatial features** & read disturbance **vulnerability**

Spatial Variation-Aware Read Disturbance Defenses:
Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

Implementation Details

Full Paper
arXiv [cs.CR] 2402.18652

Abdullah Giray Yağlıkçı

Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI ETH zürich

Svärd: Spatial Variation-Aware Read Disturbance Defenses
Integration Examples with Existing Defenses

- Integrate Svärd with five solutions

Svärd tunes **the threshold** based on the **victim row's vulnerability**

1) PARA [Kim+, ISCA'14]
- Generates **a random number**
- Compares the random number with **a threshold**
- **Refreshes the victim row** if the random number exceeds **the threshold**

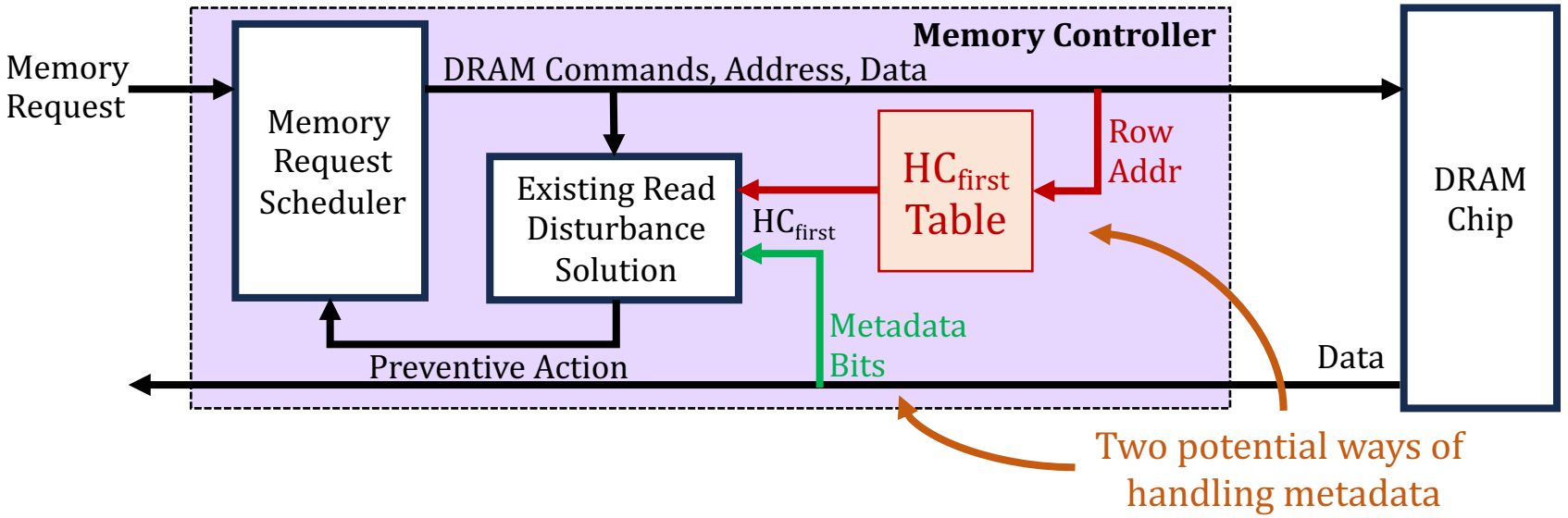
2) BlockHammer [Yaglikci+, HPCA'21]
- Counts **the number of activations** of DRAM rows
- Compares the activation count **with a threshold**
- **Throttles accesses** to the aggressor row if the activation count reaches **the threshold**

3) Hydra [Qureshi+, ISCA'22]
- Counts **the number of activations** of DRAM rows
- Compares the activation count **with a threshold**
- **Refreshes the victim row** when the activation count reaches **the threshold**

4) AQUA [Saxena+, MICRO'22] and 5) RRS [Saileshwar+, ASPLOS'22]
- Counts **the number of activations** of DRAM rows
- Compares the activation count **with a threshold**
- **Relocates the aggressor row** when the activation count reaches **the threshold**

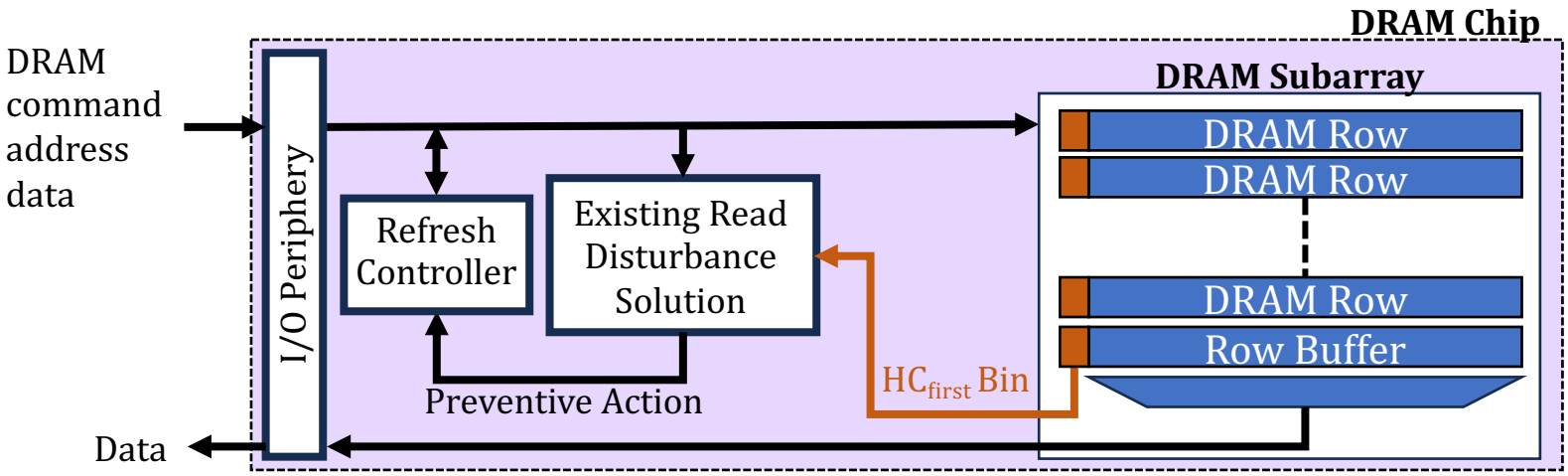
Svärd: Spatial Variation-Aware Read Disturbance Defenses
Example Implementation 1

- Classifies DRAM rows into **several vulnerability-level bins**
  - Maintains **a few (e.g., four) bits** per DRAM row
- Implemented in either the memory controller or the DRAM chip
- Two example memory controller-based implementations:



Svärd: Spatial Variation-Aware Read Disturbance Defenses
Example Implementation 2

- Classifies DRAM rows into **several vulnerability-level bins**
  - Maintains **a few (e.g., four) bits** per DRAM row
- Implemented in either the memory controller or the DRAM chip
- An example DRAM chip-based implementation:
  - Additional few (e.g., four) bits per DRAM row (e.g., 8Kb)



Spatial Variation-Aware Read Disturbance Defenses: Experimental Analysis of Real DRAM Chips and Implications on Future Solutions

Backup Slides

Abdullah Giray Yağlıkçı

Yahya Can Tuğrul Geraldo F. Oliveira İsmail Emir Yüksel

Ataberk Olgun Haocong Luo Onur Mutlu

SAFARI

ETH zürich