UNIVERSITY OF TORONTO

Learning to Trust DRAM in the Era of Worsening Rowhammer Attacks

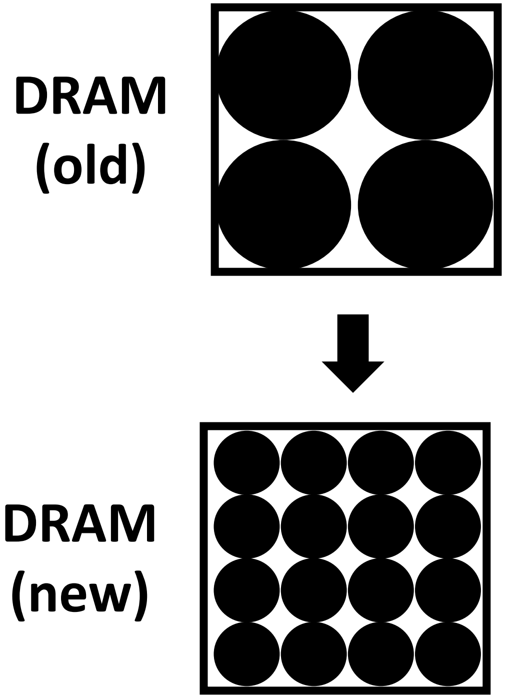
Gururaj Saileshwar

Assistant Professor, University of Toronto

4 September, 2024

Rowhammer Attacks on DRAM

DRAM Scaling for Increased Capacity



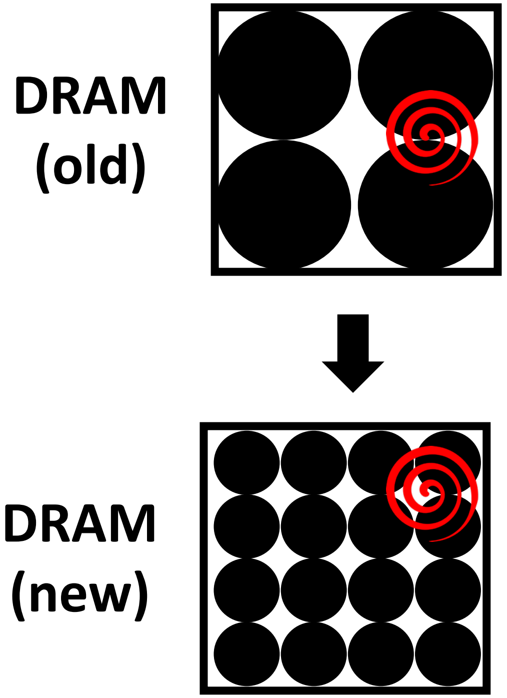
DRAM (old)

DRAM (new)

Rowhammer Attacks on DRAM

DRAM Scaling for Increased Capacity

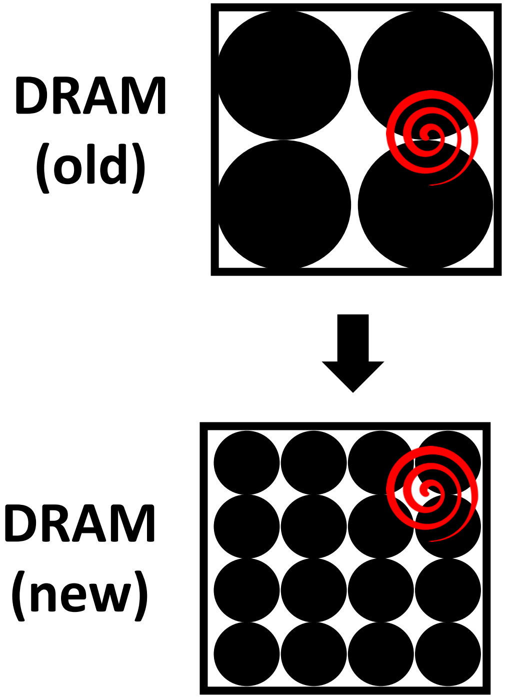
More Inter-Cell Interference



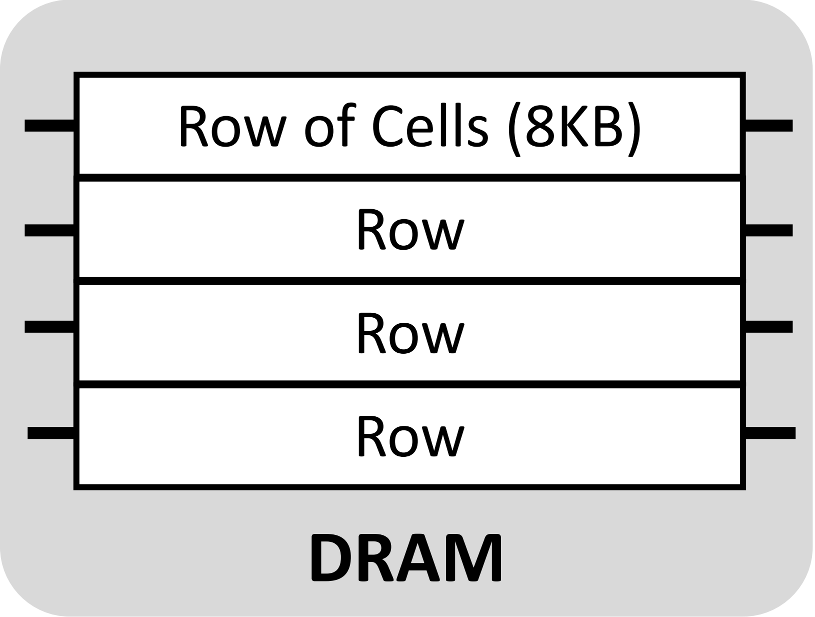
Rowhammer Attacks on DRAM

DRAM Scaling for Increased Capacity

More Inter-Cell Interference


DRAM (old)

DRAM (new)


Row of Cells (8KB)

Row

Row

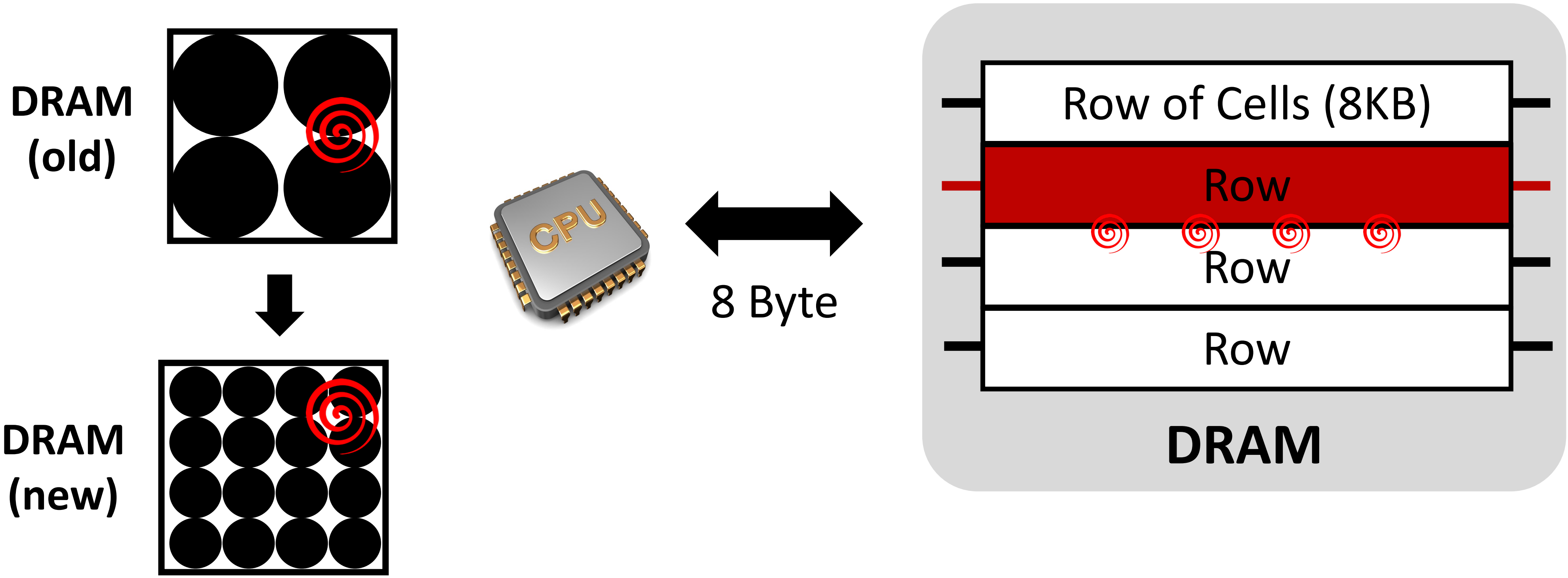
Row

DRAM

Rowhammer Attacks on DRAM

DRAM Scaling for Increased Capacity

More Inter-Cell Interference



DRAM (old)

DRAM (new)

8 Byte

Row of Cells (8KB)

Row

Row

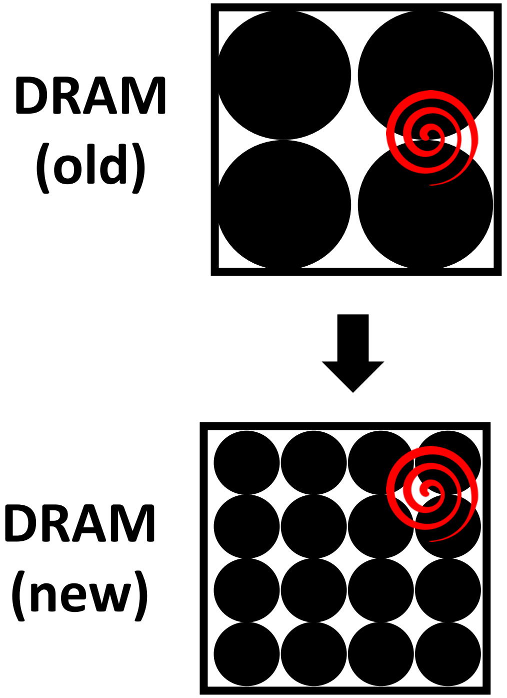
Row

DRAM

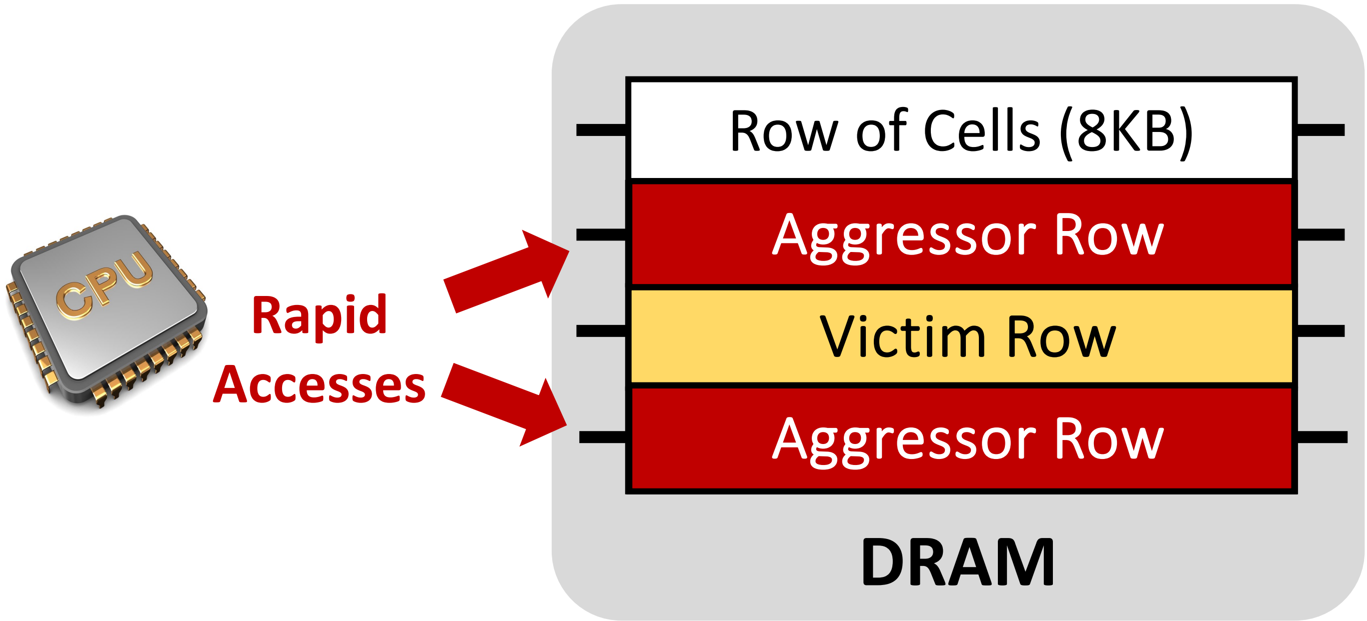
Rowhammer Attacks on DRAM

DRAM Scaling for Increased Capacity

More Inter-Cell Interference



Rowhammer Attack



Bit-Flips in Neighboring Rows

[Kim+, ISCA’14]

Rowhammer Vulnerability is Worsening

Rowhammer Threshold (Number of Activations Needed to Induce Bit-flip)
has Dropped by 30X in 8 years from 2014 to 2022

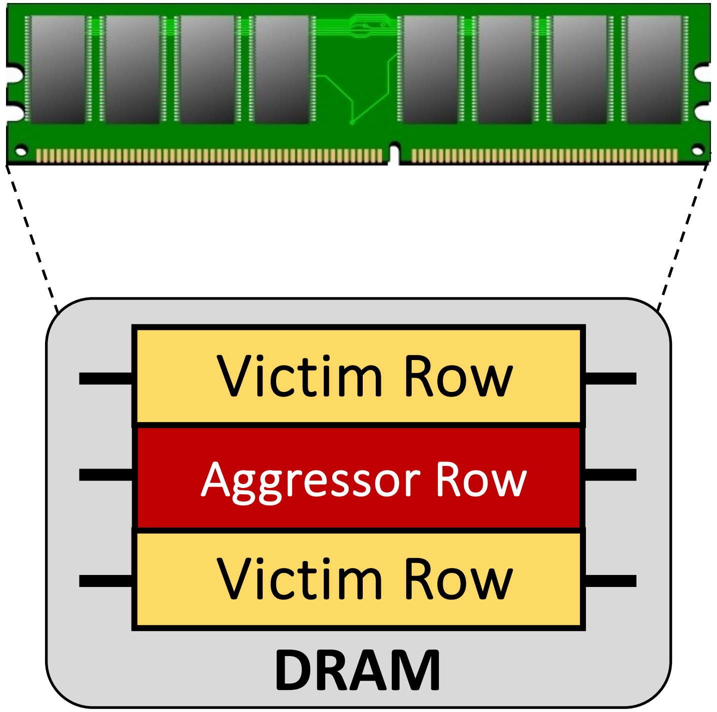
DRAM Generation	Rowhammer Threshold (TRH)
DDR3 (old)	139K [1]
DDR3 (new)	22.4K [2]
DDR4 (old)	17.5K [2]
DDR4 (new)	10K [2]
LPDDR4 (old)	16.8K [2]
LPDDR4 (new)	4.8K [2] – 9K [3]

Source: [1] - Kim+ (ISCA’14), [2] - Kim+ (ISCA’20), [3] - Kogler+ (SEC’22)

Need Defenses that are Scalable to Dropping Rowhammer Thresholds

In-DRAM Mitigation in DDR4

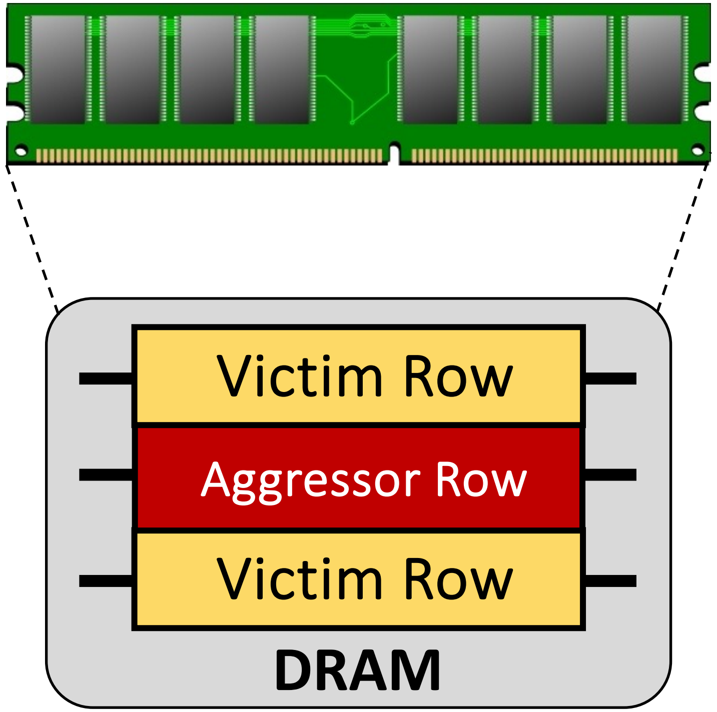
Targeted Row Refresh (TRR) in DDR4 (2015)



In-DRAM Mitigation in DDR4

Targeted Row Refresh (TRR) in DDR4 (2015)

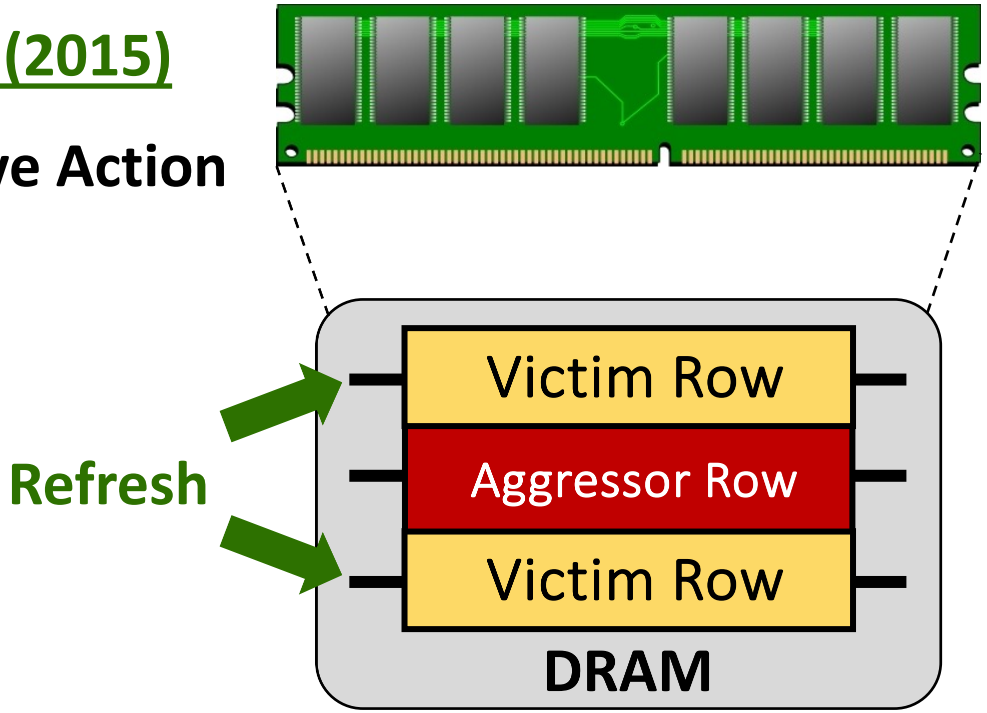
❶ Track Aggressor Rows



In-DRAM Mitigation in DDR4

Targeted Row Refresh (TRR) in DDR4 (2015)

❶ **Track Aggressor Rows** ❷ **Mitigative Action**



Refresh

Victim Row

Aggressor Row

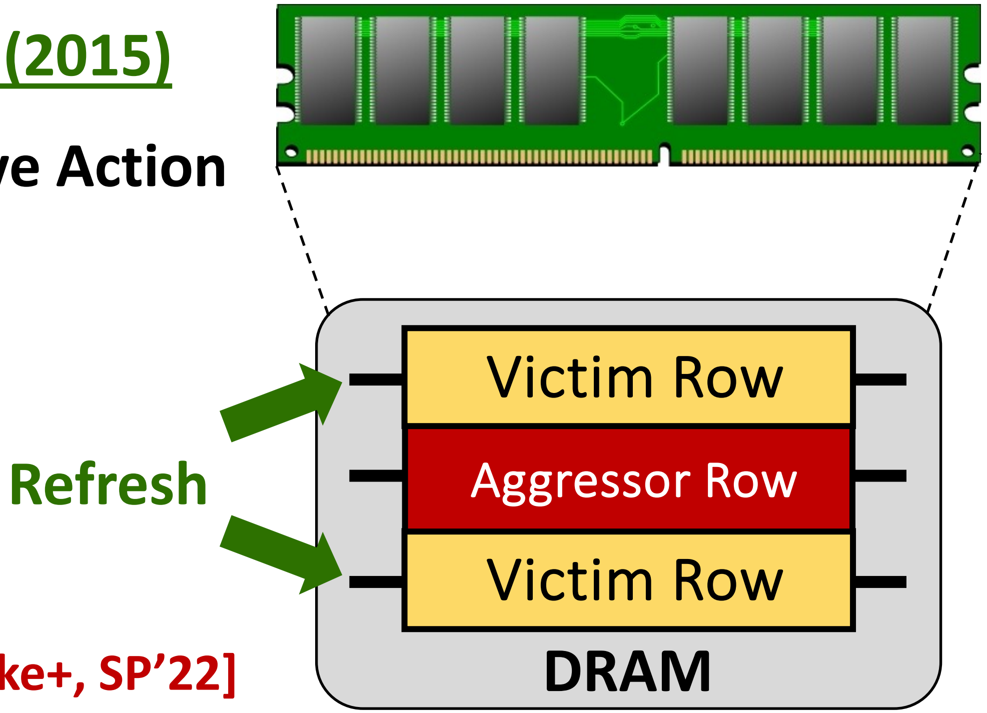
Victim Row

DRAM

Challenge-1: In-DRAM Tracking Solutions Broken!

Targeted Row Refresh (TRR) in DDR4 (2015)

❶ ~~Track Aggressor Rows~~ ❷ Mitigative Action

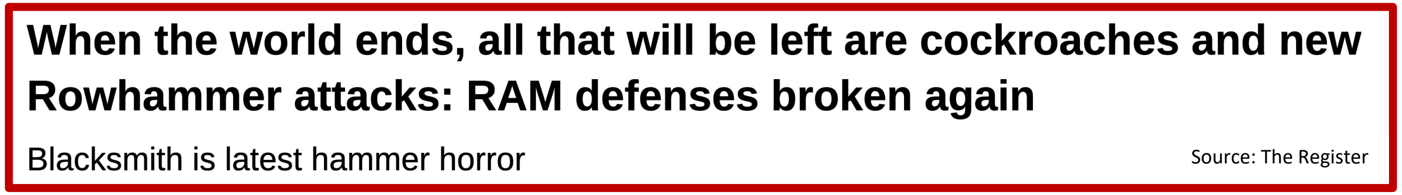


TRResspass Breaks TRR Tracker [Frigo+, SP’20]

Poor Rowhammer Fixes On DDR4 DRAM Chips Re-Enable Bit Flipping Attacks

Source: The Hacker News

Blacksmith Attack: All DDR4 DRAM Vulnerable [Jattke+, SP’22]



When the world ends, all that will be left are cockroaches and new Rowhammer attacks: RAM defenses broken again

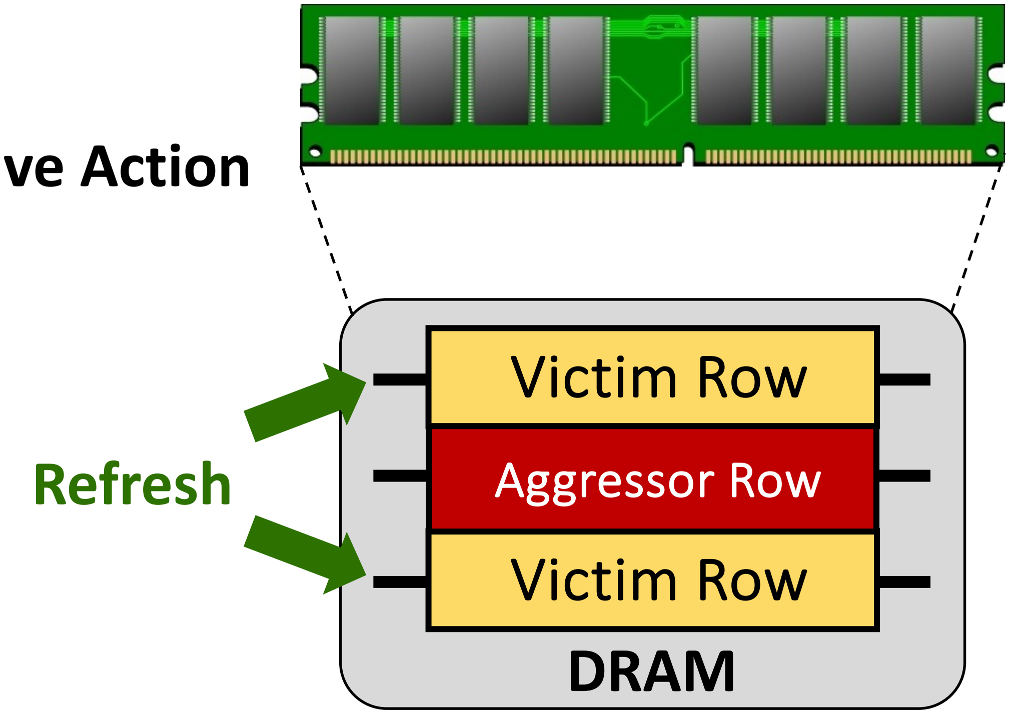
Blacksmith is latest hammer horror

Source: The Register

Challenge-1: In-DRAM Tracking Solutions Broken!

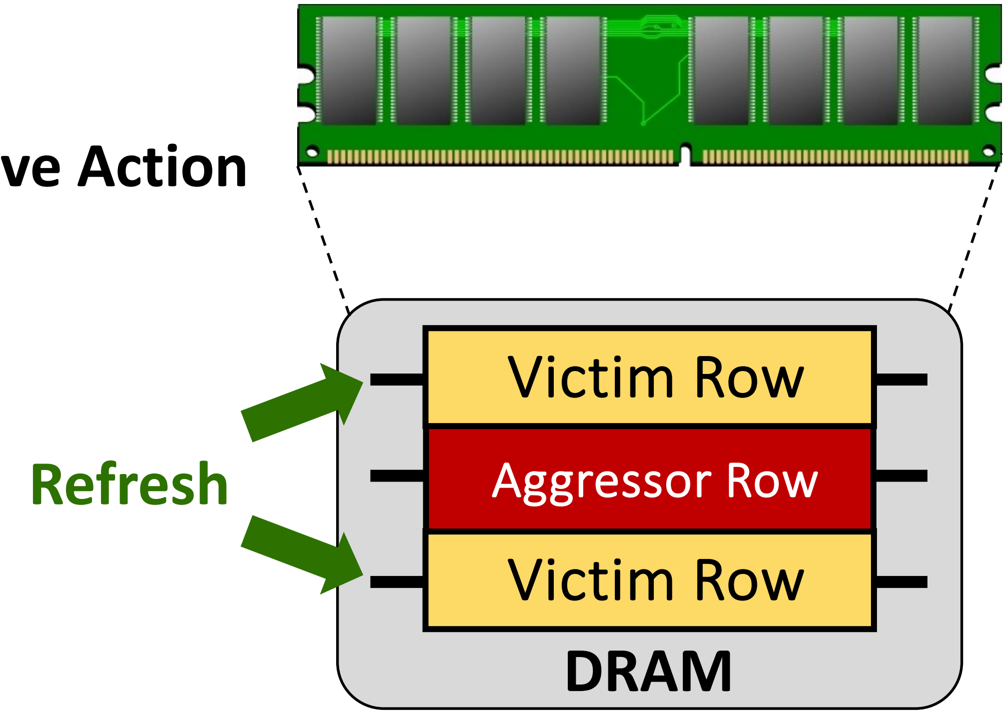
❶ **Track Aggressor Rows** ❷ **Mitigative Action**

?

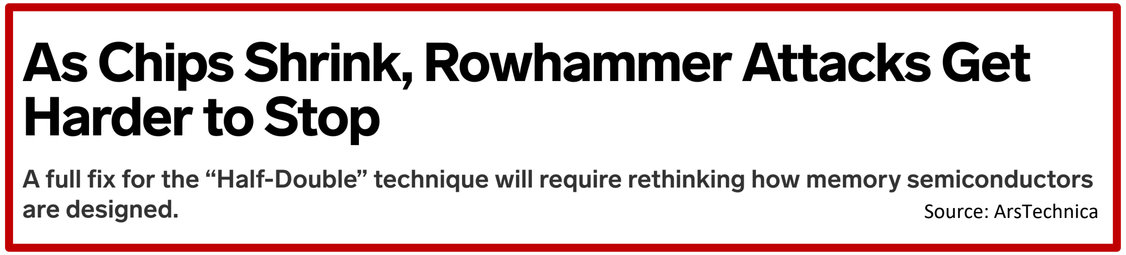


Challenge-2: New Attacks on Victim-Focused Mitigation

❶ Track Aggressor Rows ❷ Mitigative Action



Google's Half-Double Attack:
Exploits Mitigative Refresh [SEC'22]



As Chips Shrink, Rowhammer Attacks Get Harder to Stop

A full fix for the "Half-Double" technique will require rethinking how memory semiconductors are designed.

Source: ArsTechnica

Challenge-2: New Attacks on Victim-Focused Mitigation

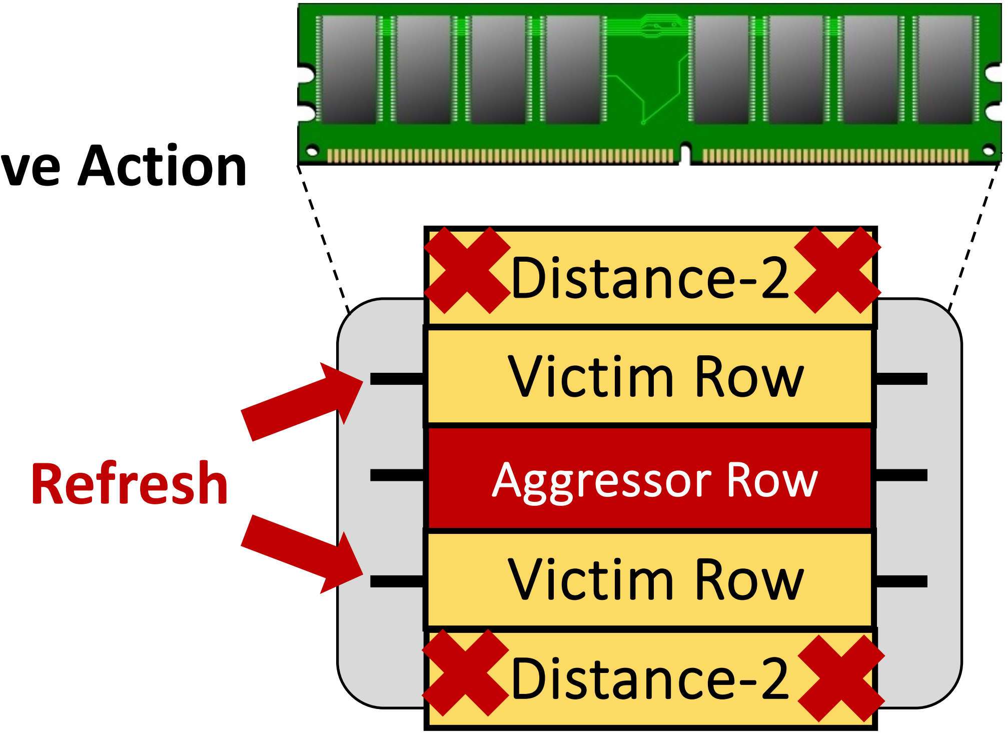
❶ Track Aggressor Rows ❷ Mitigative Action

Google's Half-Double Attack:
Exploits Mitigative Refresh [SEC'22]

As Chips Shrink, Rowhammer Attacks Get Harder to Stop

A full fix for the "Half-Double" technique will require rethinking how memory semiconductors are designed.

Source: ArsTechnica



Distance-2

Victim Row

Aggressor Row

Victim Row

Distance-2

Refresh

Need New Mitigative Actions to Mitigate Rowhammer

Our Scalable & Practical Defenses Against Rowhammer

Challenge: New Attacks on Victim-Focused Mitigations in-DRAM?

New Aggressor Focused Mitigation

ASPLOS'22: Randomized Row-Swap

HPCA'23: Scalable & Secure Row-Swap

🏆 Best Paper Award

Challenge: In-DRAM Tracking Solutions Broken?

Secure Tracking Solutions in-DRAM

ISCA'24: PrIDE - Probabilistic In-DRAM Tracker

Scalable to sub-500 TRH

Defense-in-Depth

DSN'23: PTGuard – Integrity Protection for Targets of Rowhammer (Page-Tables)

Agenda

Introduction

New Mitigative Actions for Rowhammer

Randomized Row-Swaps [ASPLOS 2022, HPCA 2023]

Secure In-DRAM Tracking

PrIDE: Probabilistic In-DRAM Tracker [ISCA 2024]

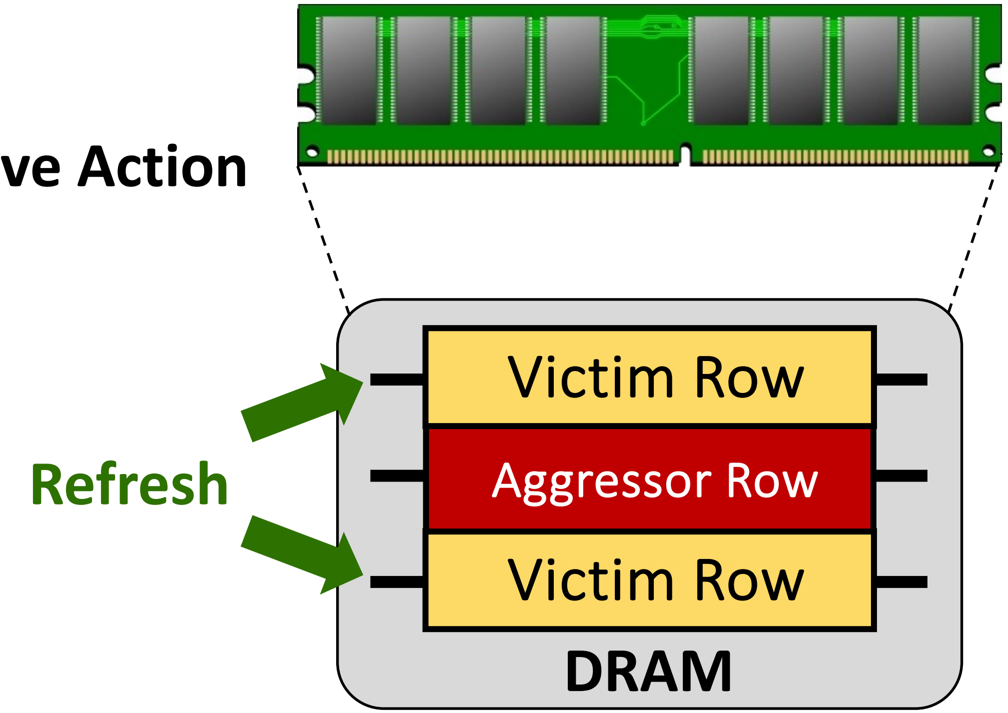
Defense in Depth Solutions

PT-Guard [DSN 2023]

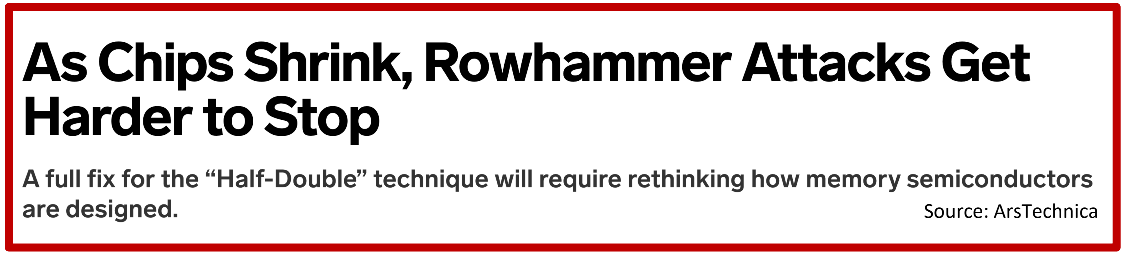
Conclusion

Motivation: Attacks On Victim-Focused Mitigation

☑ Track Aggressor Rows ❷ Mitigative Action



Google's Half-Double Attack:
Exploits Mitigative Refresh [2021, SEC'22]



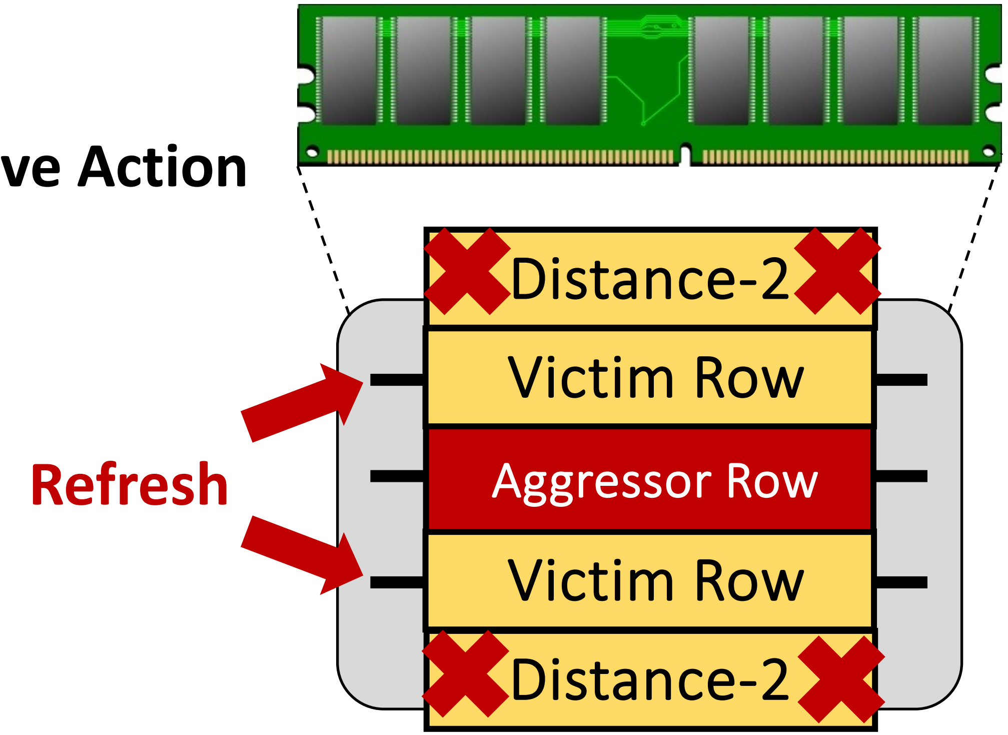
As Chips Shrink, Rowhammer Attacks Get Harder to Stop

A full fix for the "Half-Double" technique will require rethinking how memory semiconductors are designed.

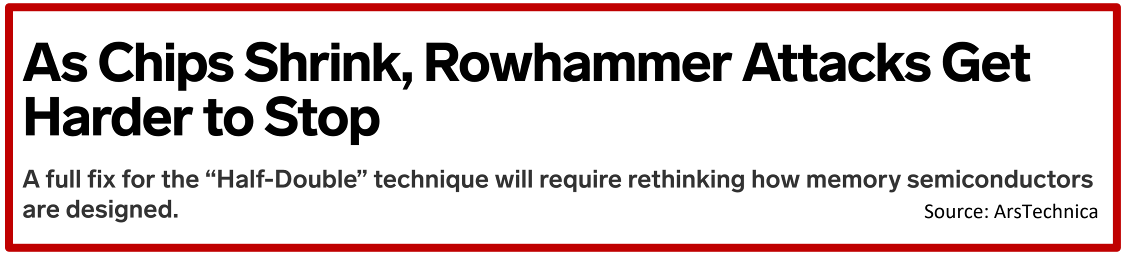
Source: ArsTechnica

Motivation: Attacks On Victim-Focused Mitigation

☑ Track Aggressor Rows ❷ Mitigative Action



Google's Half-Double Attack:
Exploits Mitigative Refresh [2021, SEC'22]



As Chips Shrink, Rowhammer Attacks Get Harder to Stop

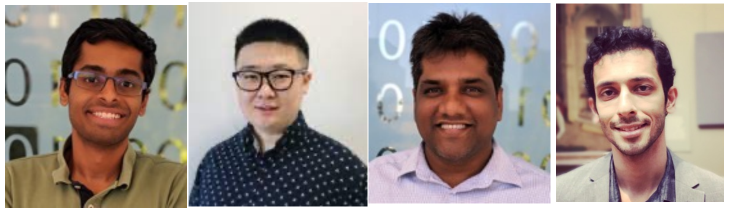
A full fix for the "Half-Double" technique will require rethinking how memory semiconductors are designed.

Source: ArsTechnica

Need New Mitigative Action Resilient to New Attack Patterns
(without requiring knowledge of DRAM mapping function)

Randomized Row-Swap: Mitigating Row Hammer By Breaking Spatial Correlation Between Aggressor and Victim Rows

ASPLOS 2022, Lausanne, Switzerland



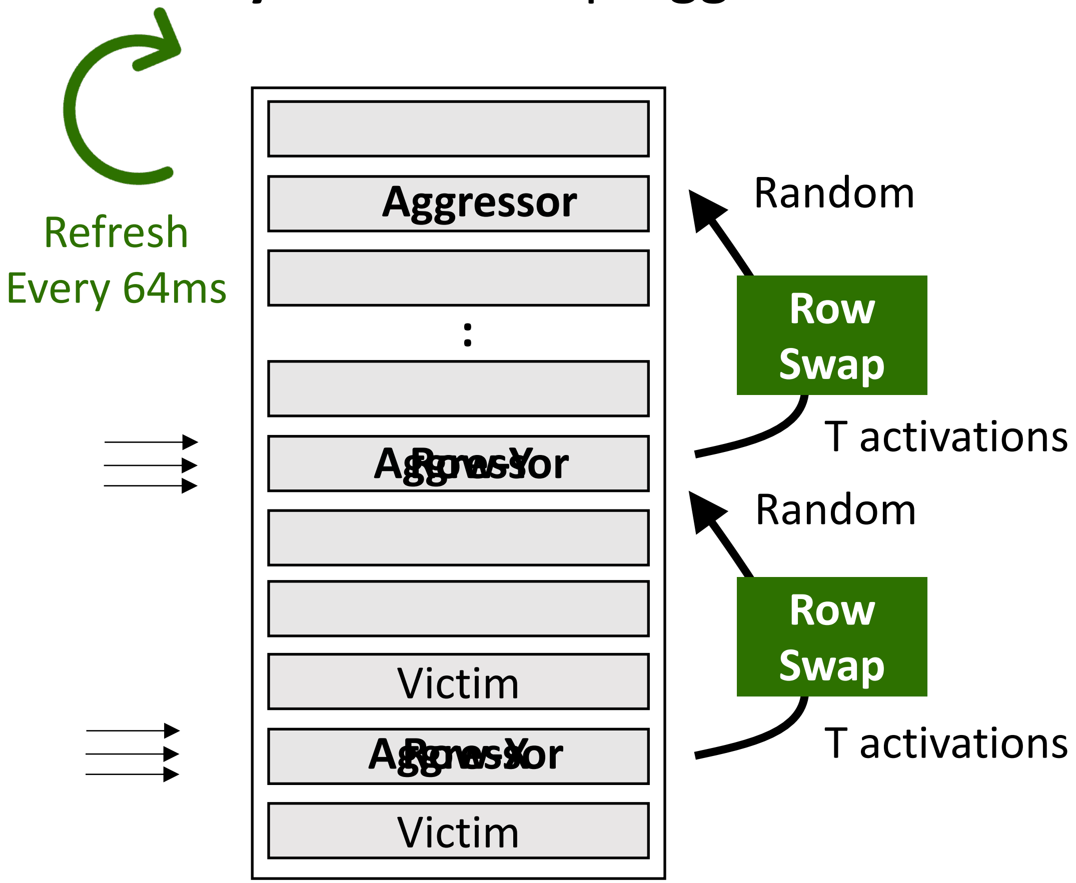
Gururaj Saileshwar, Bolin Wang, Moinuddin Qureshi, Prashant Nair

GT Georgia Tech

UBC THE UNIVERSITY OF BRITISH COLUMBIA

Aggressor Focused Mitigation: Randomized Row-Swap

Key Idea: Remap Aggressor Rows to Break Spatial Correlation with Victim Rows



Security Guarantee: No Row Crosses Rowhammer Threshold (TRH = 4800) Activations within 64ms

Lower T (Swap Threshold) → Better Security

Swap Every T Activations	Attack Time
T = TRH/5	6.9 days
T = TRH/6	**3.8 years**
T = TRH/7	762 years

Security Analysis

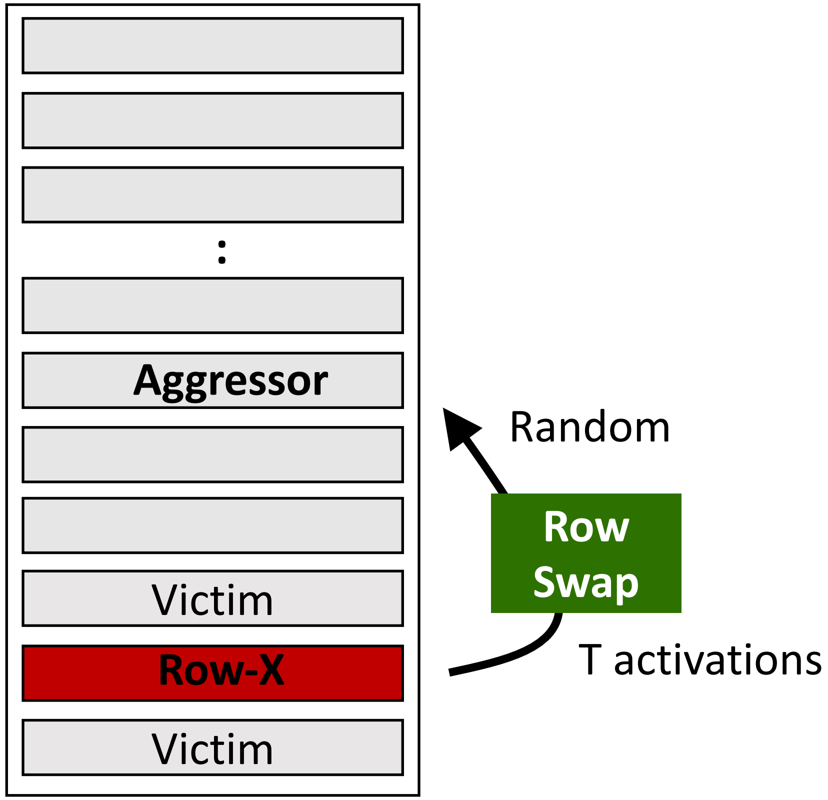
TRH=4800 → Minimum Activations in 64ms on Row for Rowhammer via Any Pattern

(Single-sided, Double-Sided, Half-Double)

Security Analysis

TRH=4800 → Minimum Activations in 64ms on Row for Rowhammer via Any Pattern

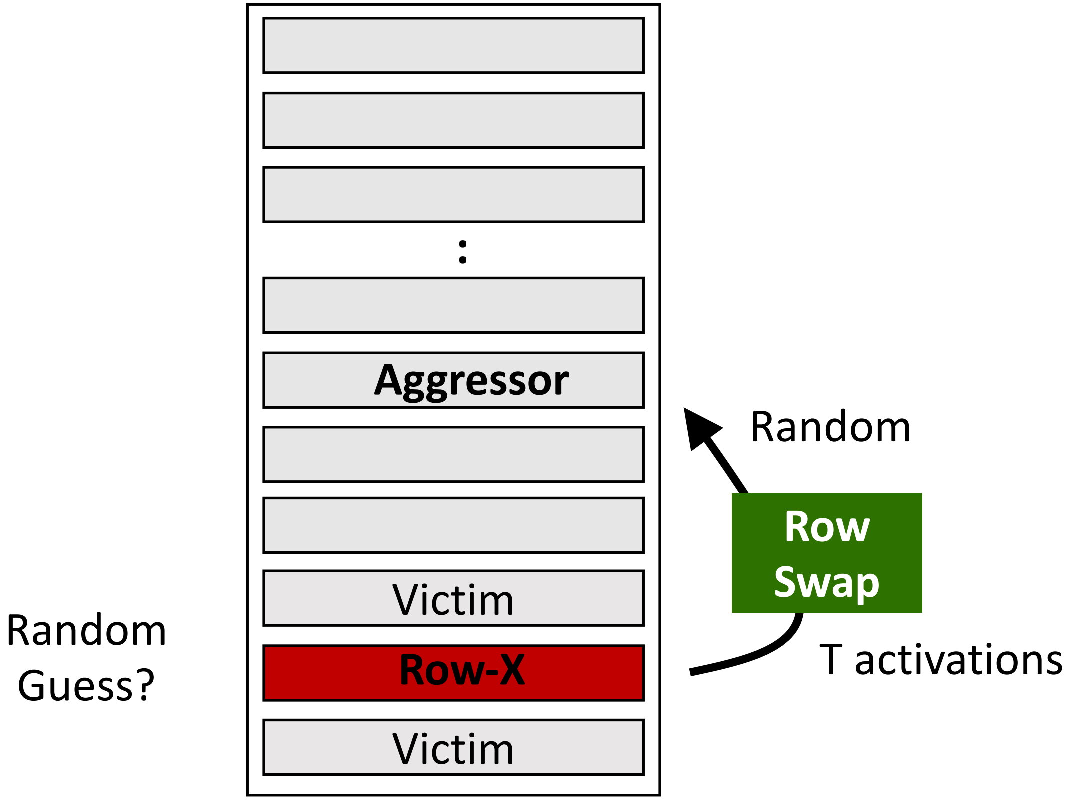
(Single-sided, Double-Sided, Half-Double)



Security Analysis

TRH=4800 → Minimum Activations in 64ms on Row for Rowhammer via Any Pattern

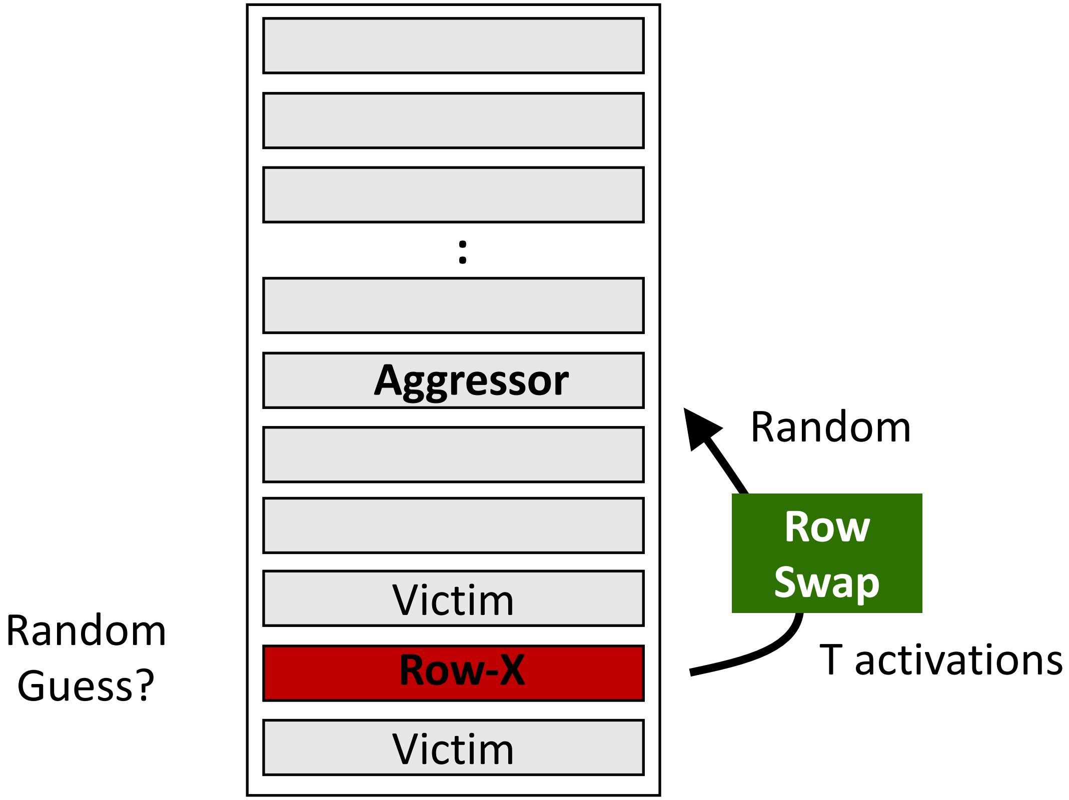
(Single-sided, Double-Sided, Half-Double)



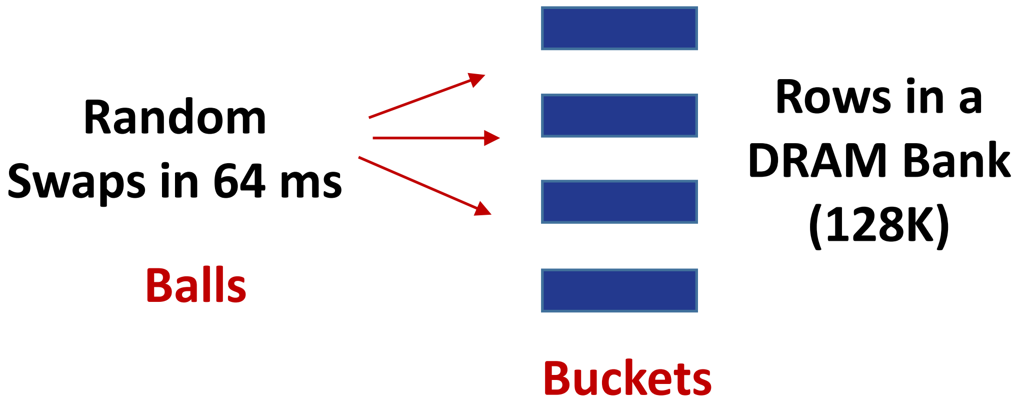
Security Analysis

TRH=4800 → Minimum Activations in 64ms on Row for Rowhammer via Any Pattern

(Single-sided, Double-Sided, Half-Double)



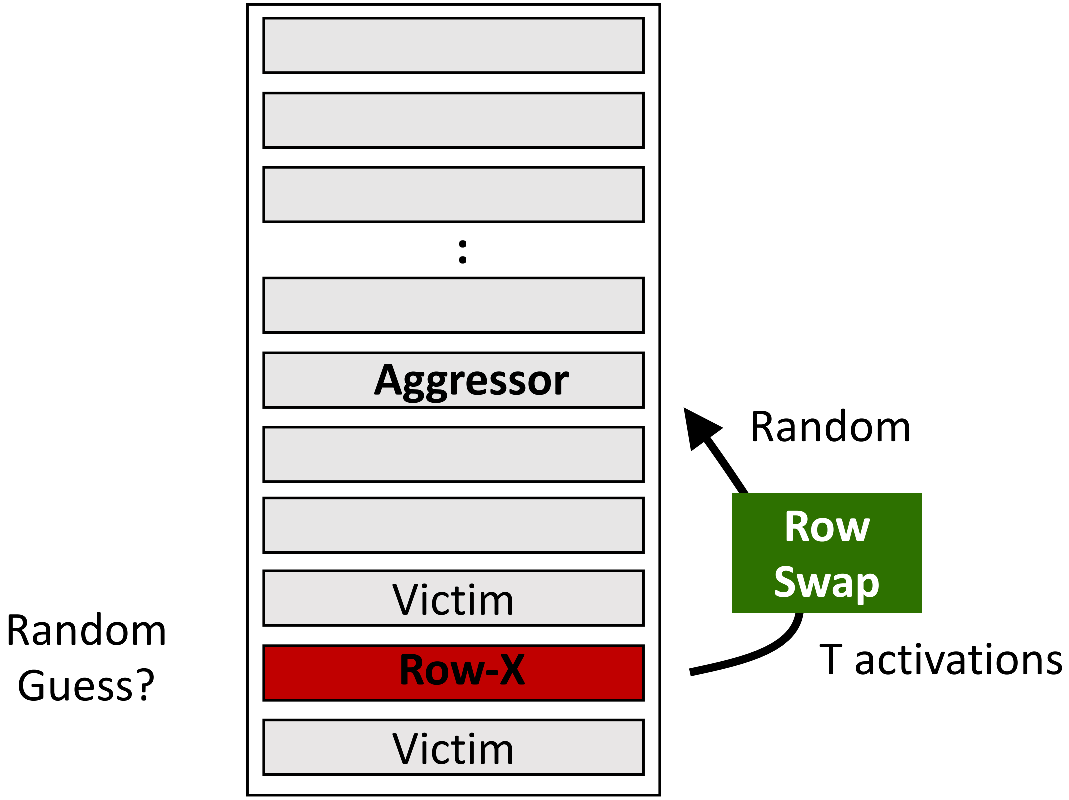
Buckets and Balls Problem



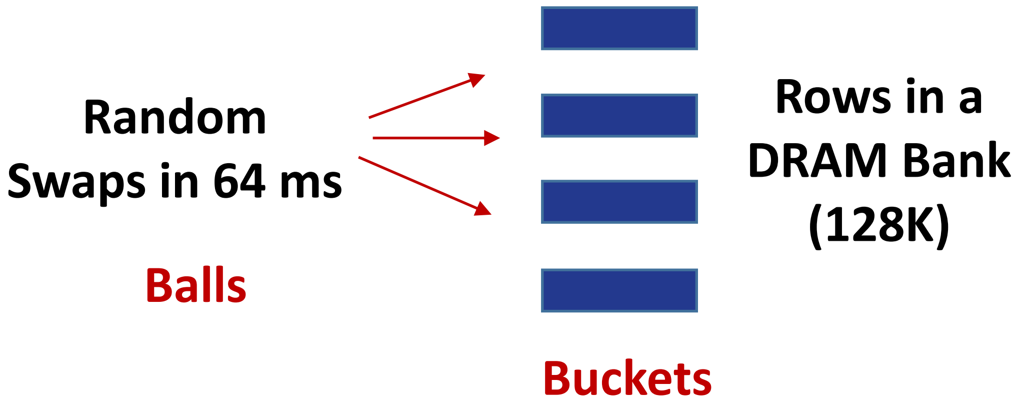
Security Analysis

TRH=4800 → Minimum Activations in 64ms on Row for Rowhammer via Any Pattern

(Single-sided, Double-Sided, Half-Double)

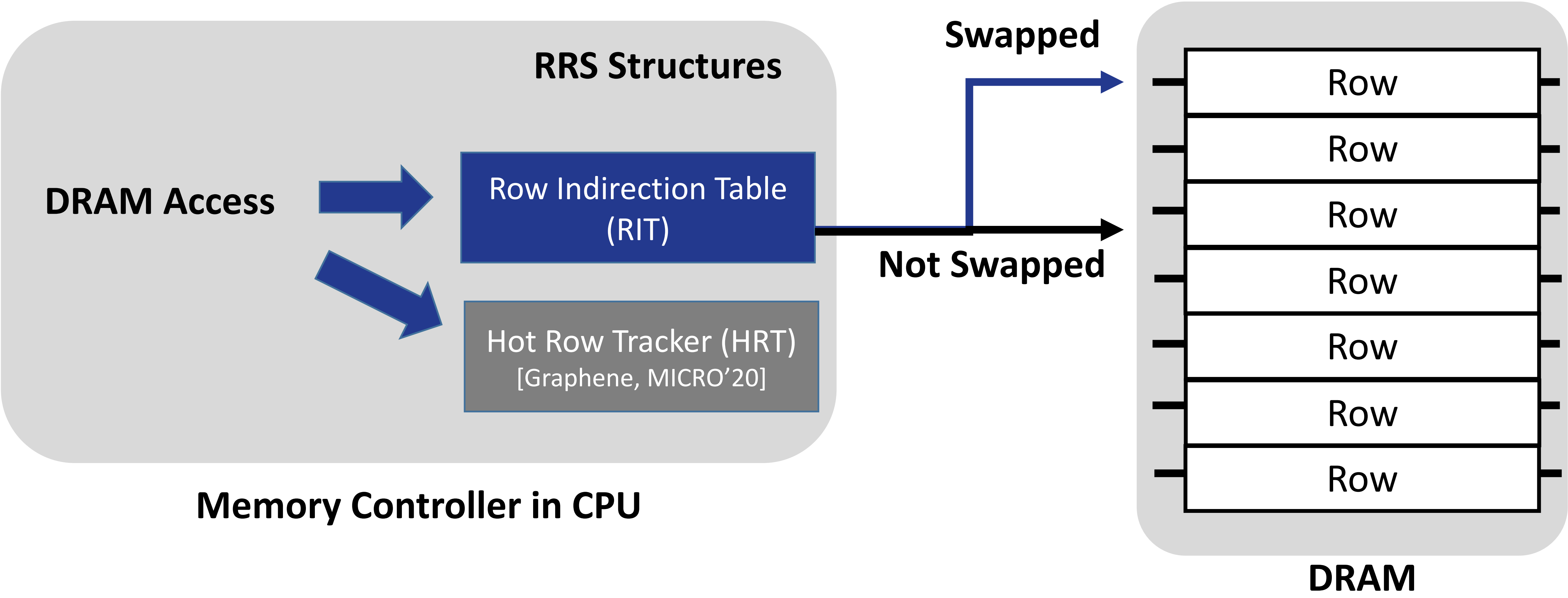


Buckets and Balls Problem

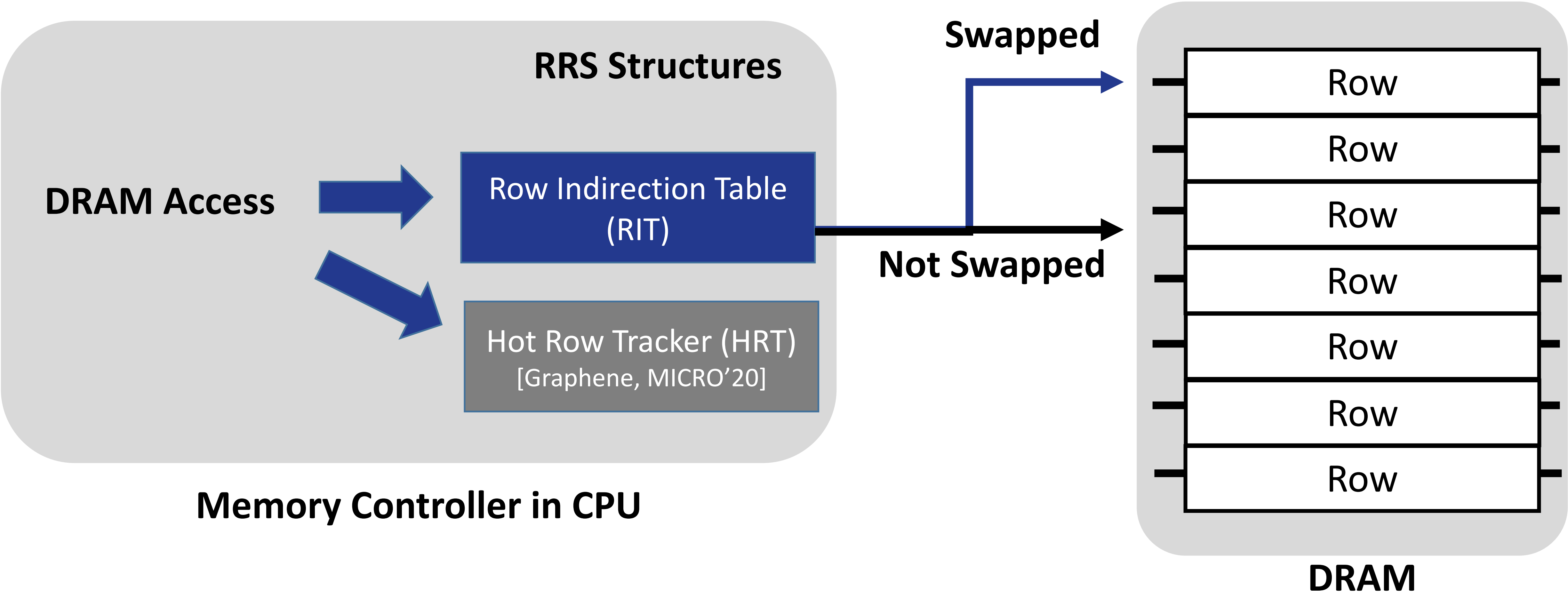


Swap Every T Activations	Attack Time
T = TRH/5	6.9 days
T = TRH/6	**3.8 years**
T = TRH/7	762 years

Implementation of Randomized Row Swap



Implementation of Randomized Row Swap



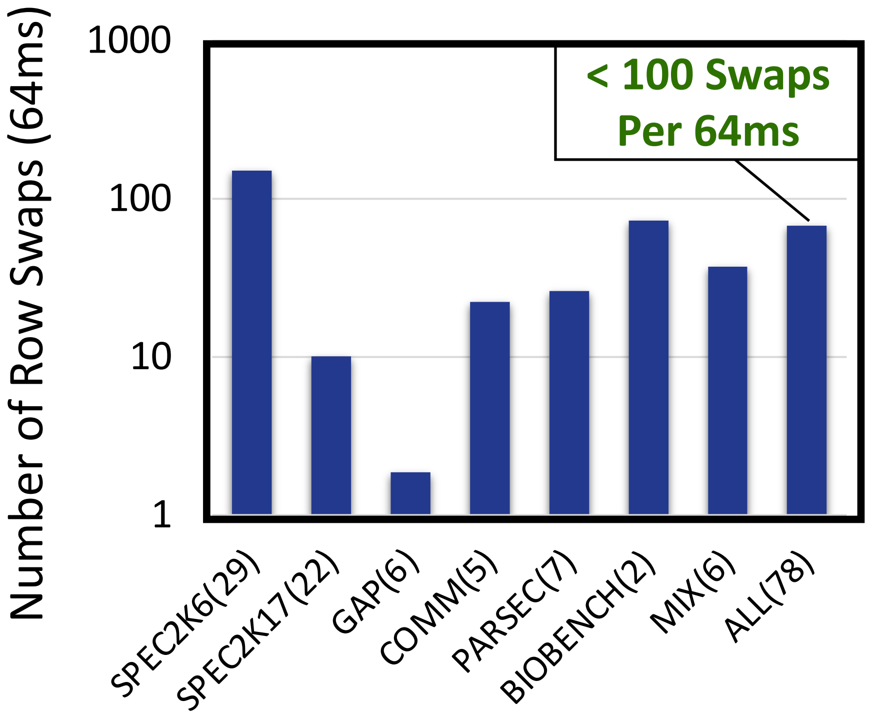
RIT Stores Tuples of Swapped Rows → RIT + HRT = 45 KB Per DRAM Bank → 700KB Per Rank

Performance Impact of Row Swaps

Config: 8-core OOO, 16GB DRAM (1 Rank). Rowhammer Threshold of 4.8K.

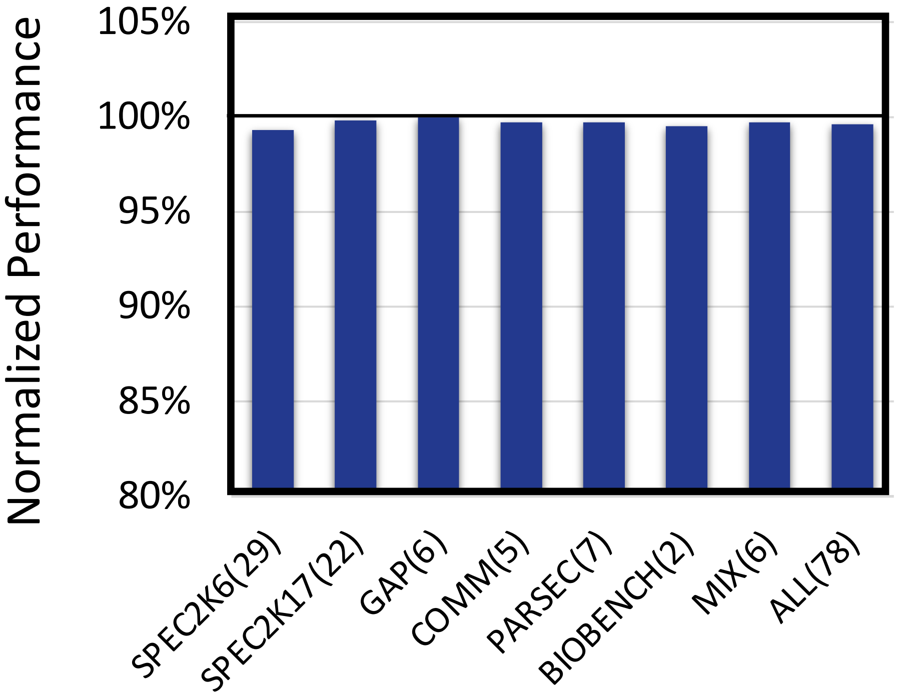
Frequency of Row Swaps Per 64ms

(1.5 microseconds per swap)



Negligible Performance Impact

(0.4% slowdown on average)

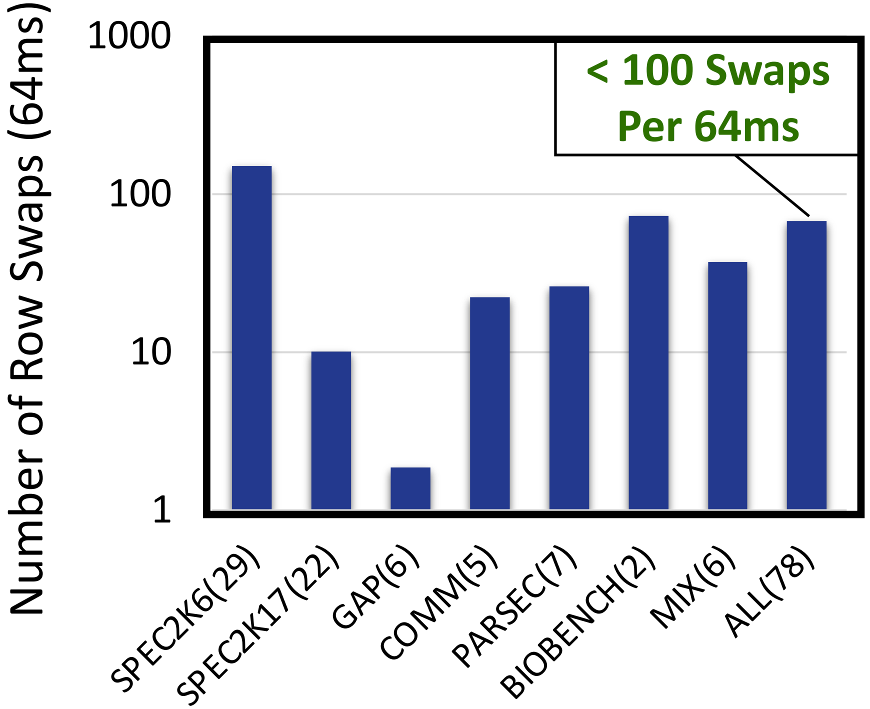


Performance Impact of Row Swaps

Config: 8-core OOO, 16GB DRAM (1 Rank). Rowhammer Threshold of 4.8K.

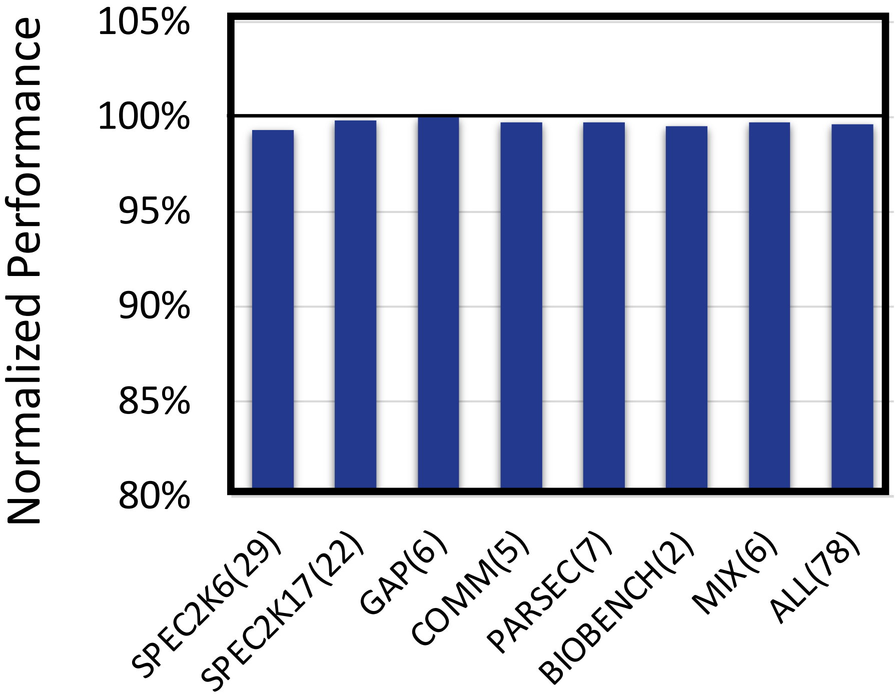
Frequency of Row Swaps Per 64ms

(1.5 microseconds per swap)



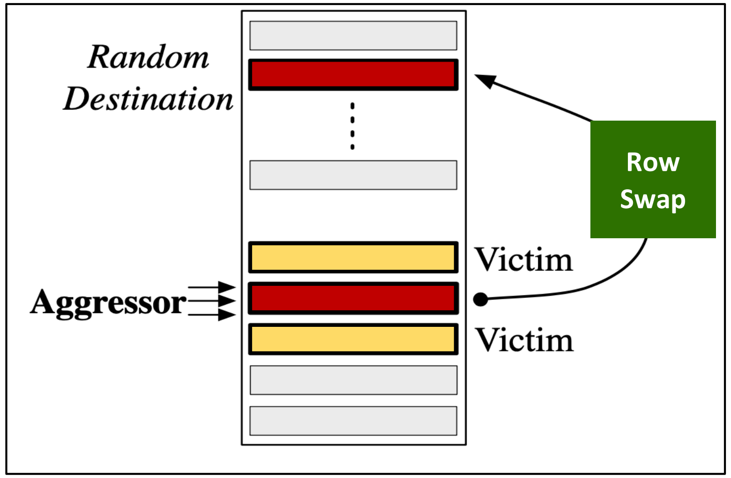
Negligible Performance Impact

(0.4% slowdown on average)



Randomized Row Swap has negligible performance impact due to infrequent swaps

Takeways from Randomized Row Swap



New Aggressor-Focused Mitigation
CPU-side Implementation, compatible with commodity DRAM

Incurs Modest Costs at TRH of 4.8K (0.4% slowdown, 45KB SRAM/bank)

Scalable and Secure Row-Swap: Efficient and Safe Rowhammer Mitigation in Memory Systems

HPCA 2023, Montreal, Canada

Best Paper Award

Jeonghyun Woo, Gururaj Saileshwar, Prashant Nair

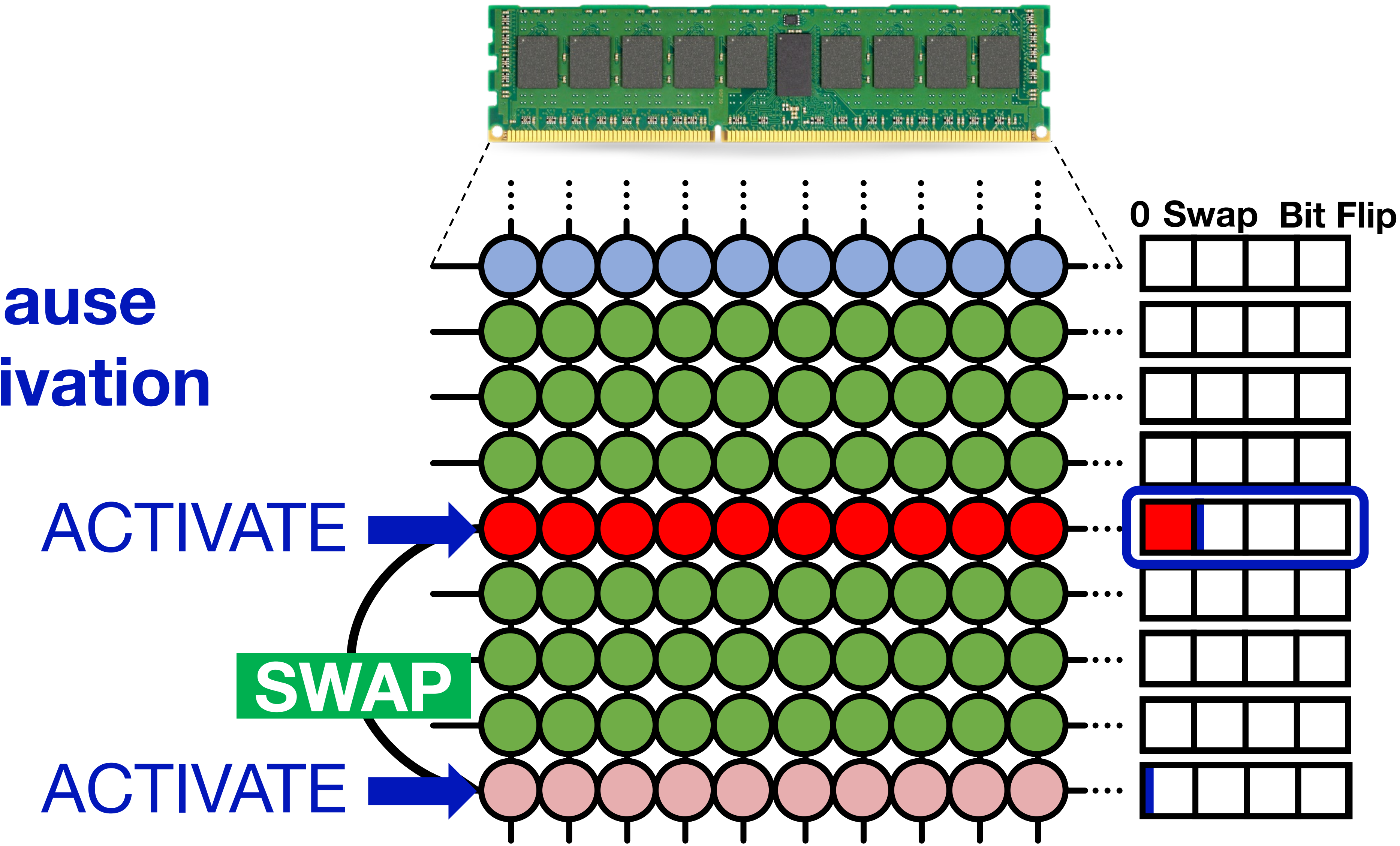
NVIDIA *

STAR LAB

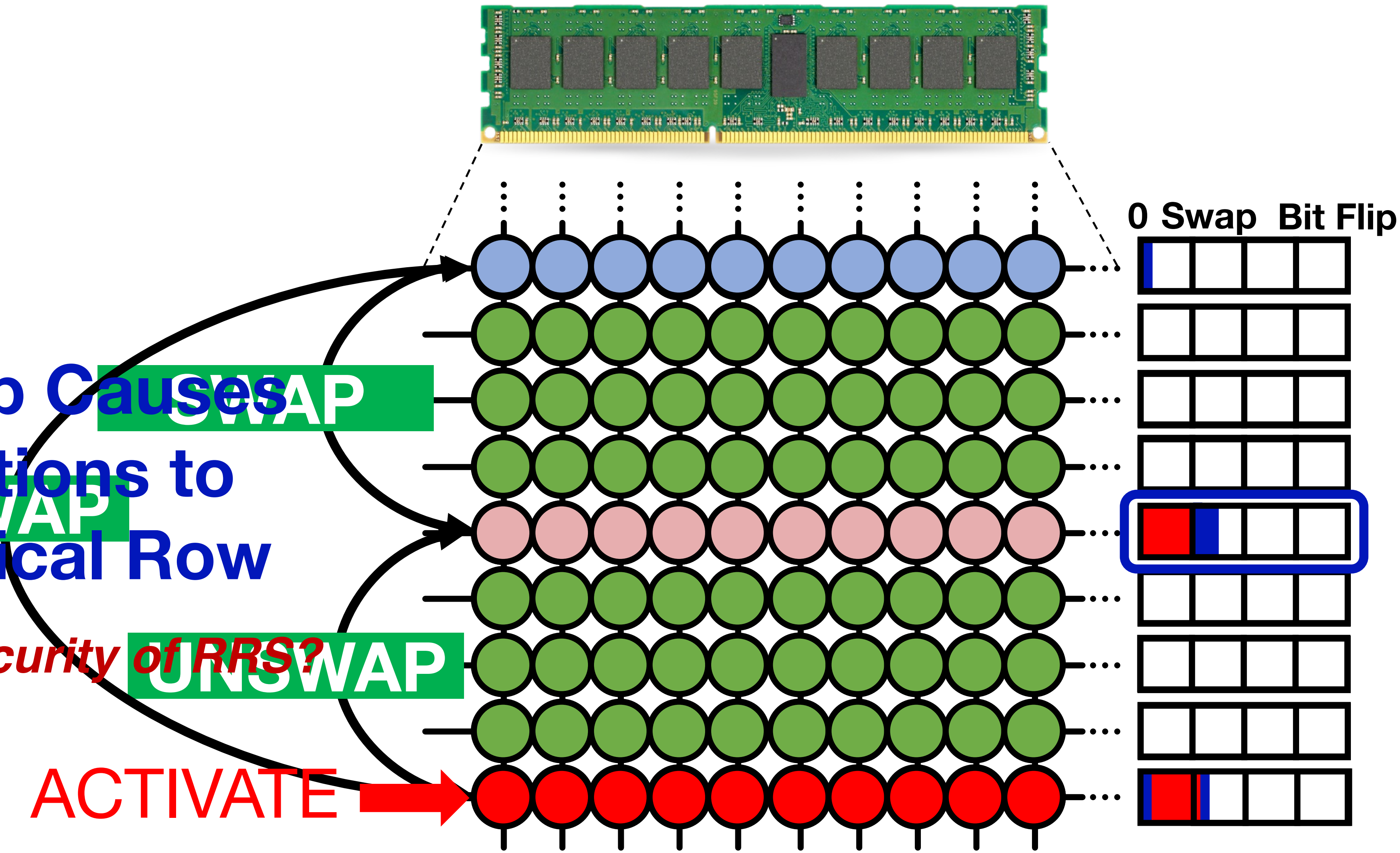
UBC THE UNIVERSITY OF BRITISH COLUMBIA

RRS Security Pitfall: Latent Activations

Swaps Cause Latent Activation



RRS Security Pitfall: Latent Activations



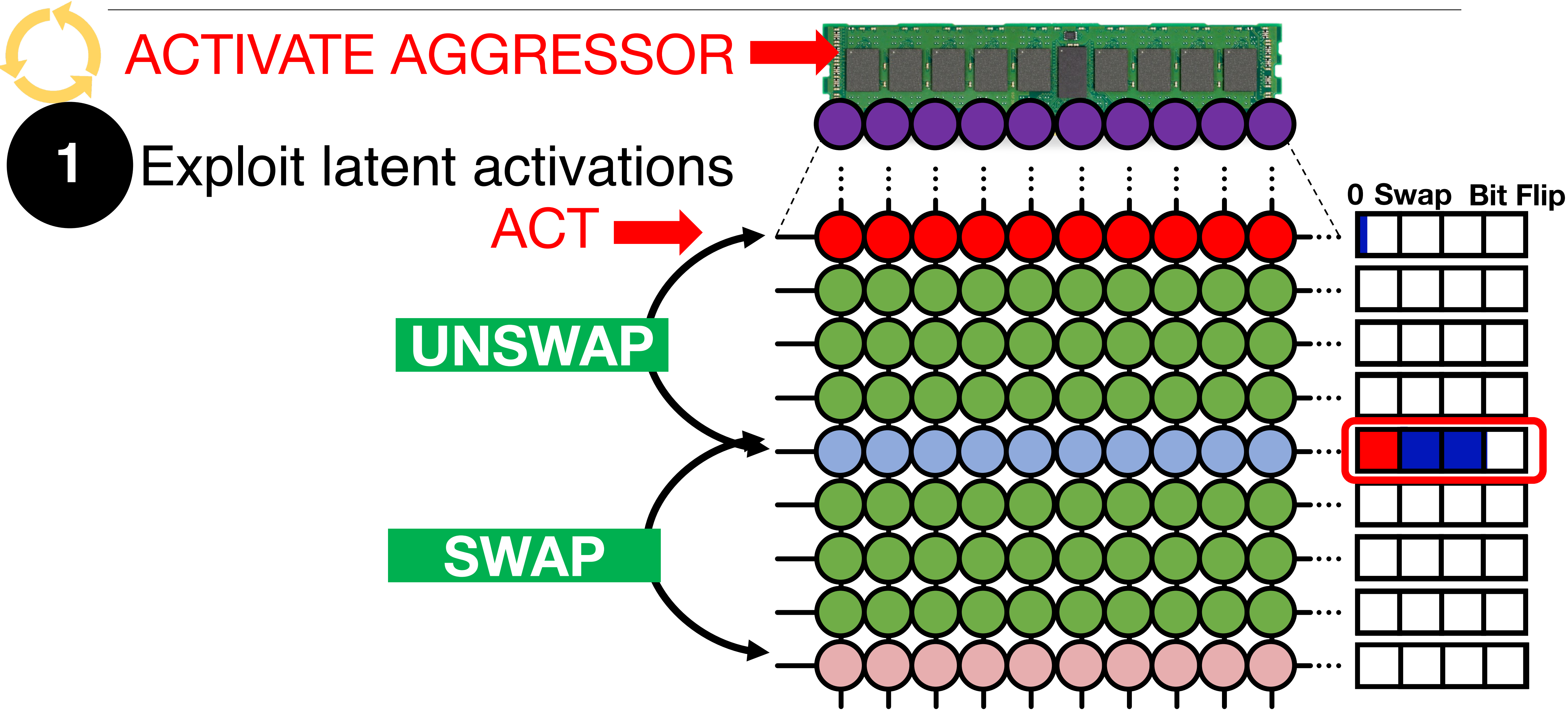
UnSwap-Swap Causes Extra Activations to Original Physical Row

Can this break the security of RRS?

Juggernaut Attack

ACTIVATE AGGRESSOR

1 Exploit latent activations



Juggernaut Attack

Random Guess

1 Exploit latent activations

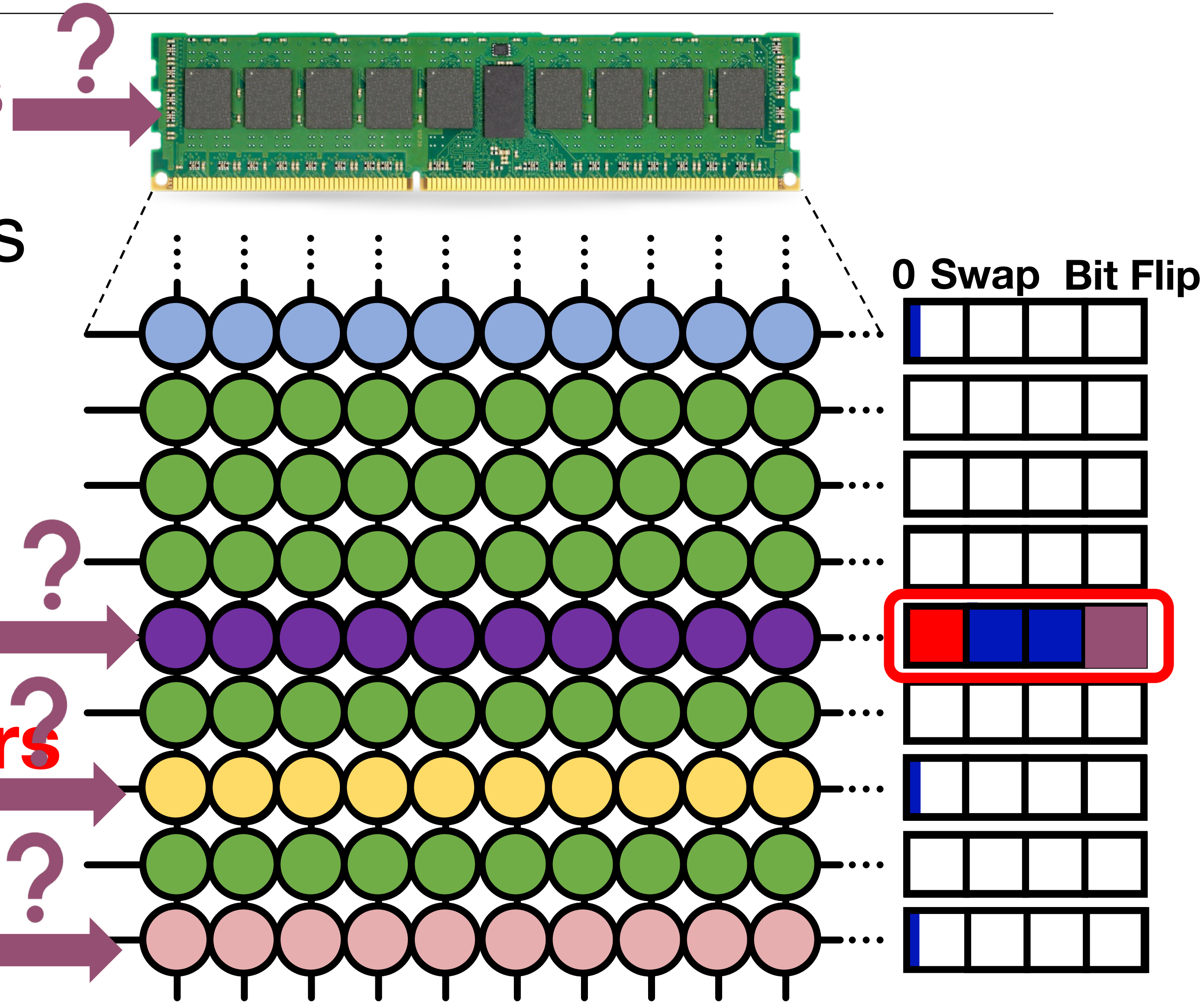
2 Guess Biased Row

Random Guess

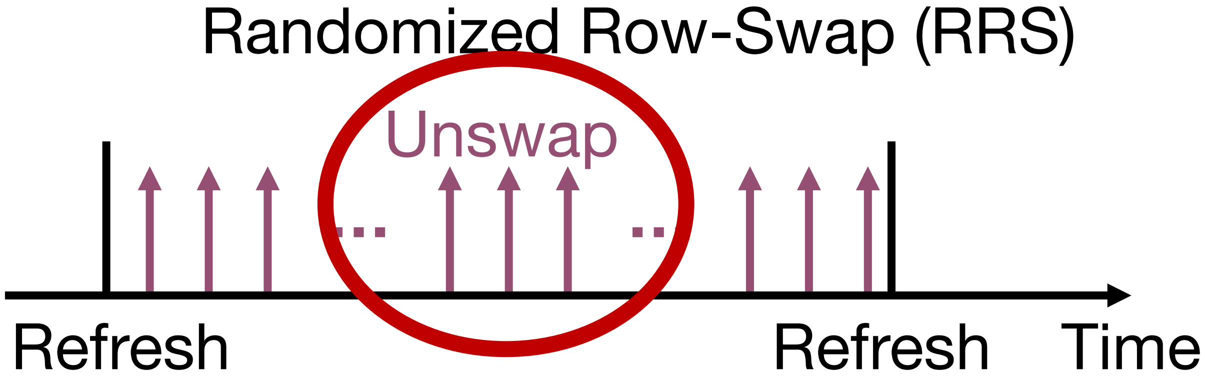
Break RRS in < 4 hours

Random Guess

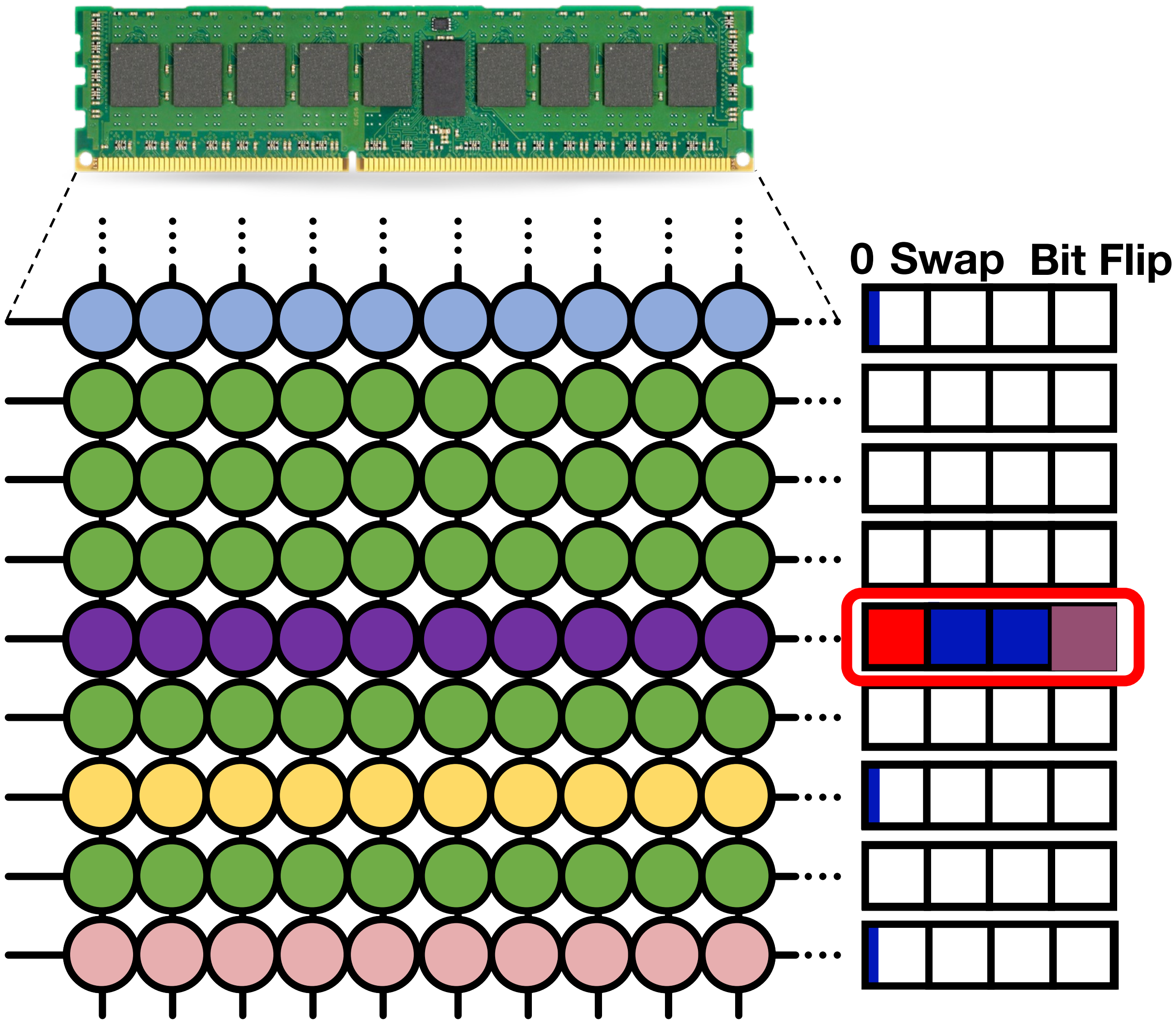
Random Guess



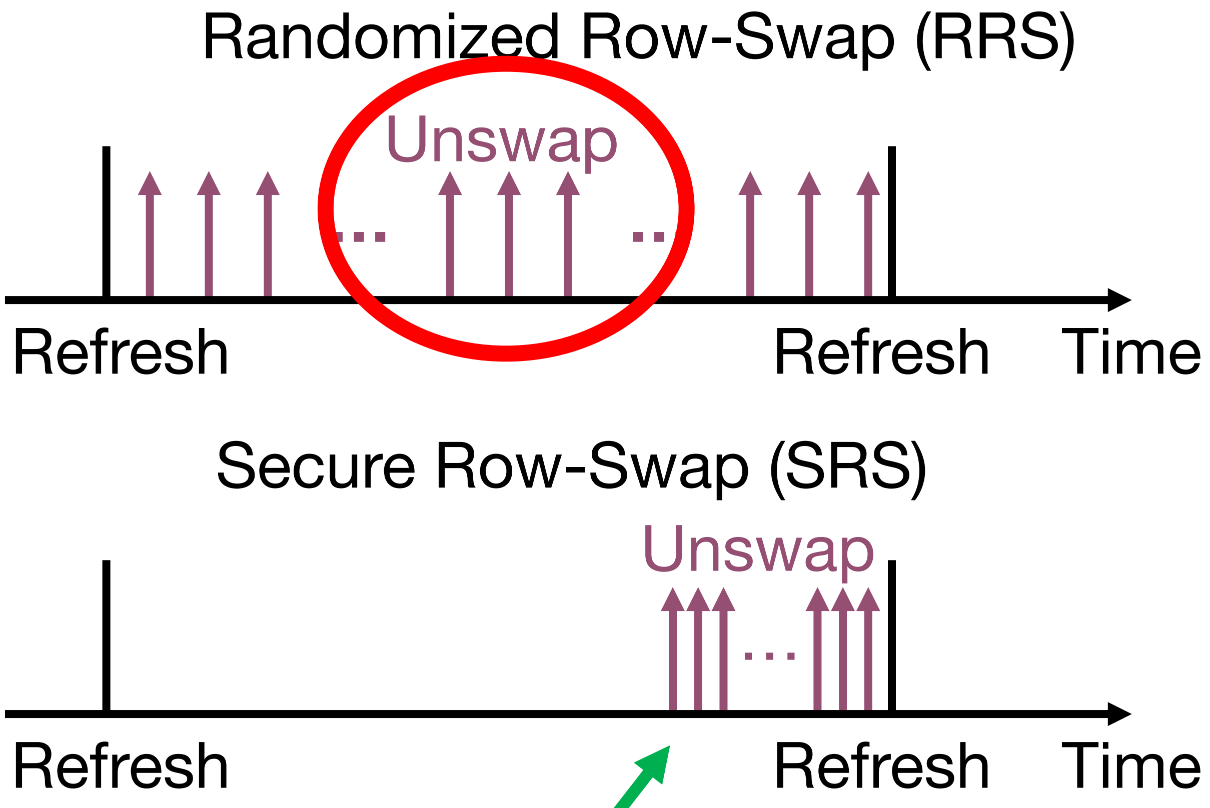
RRS Suffers Vulnerability Due to Unswaps



Unswaps in RRS Required Due to Tracking Complexity



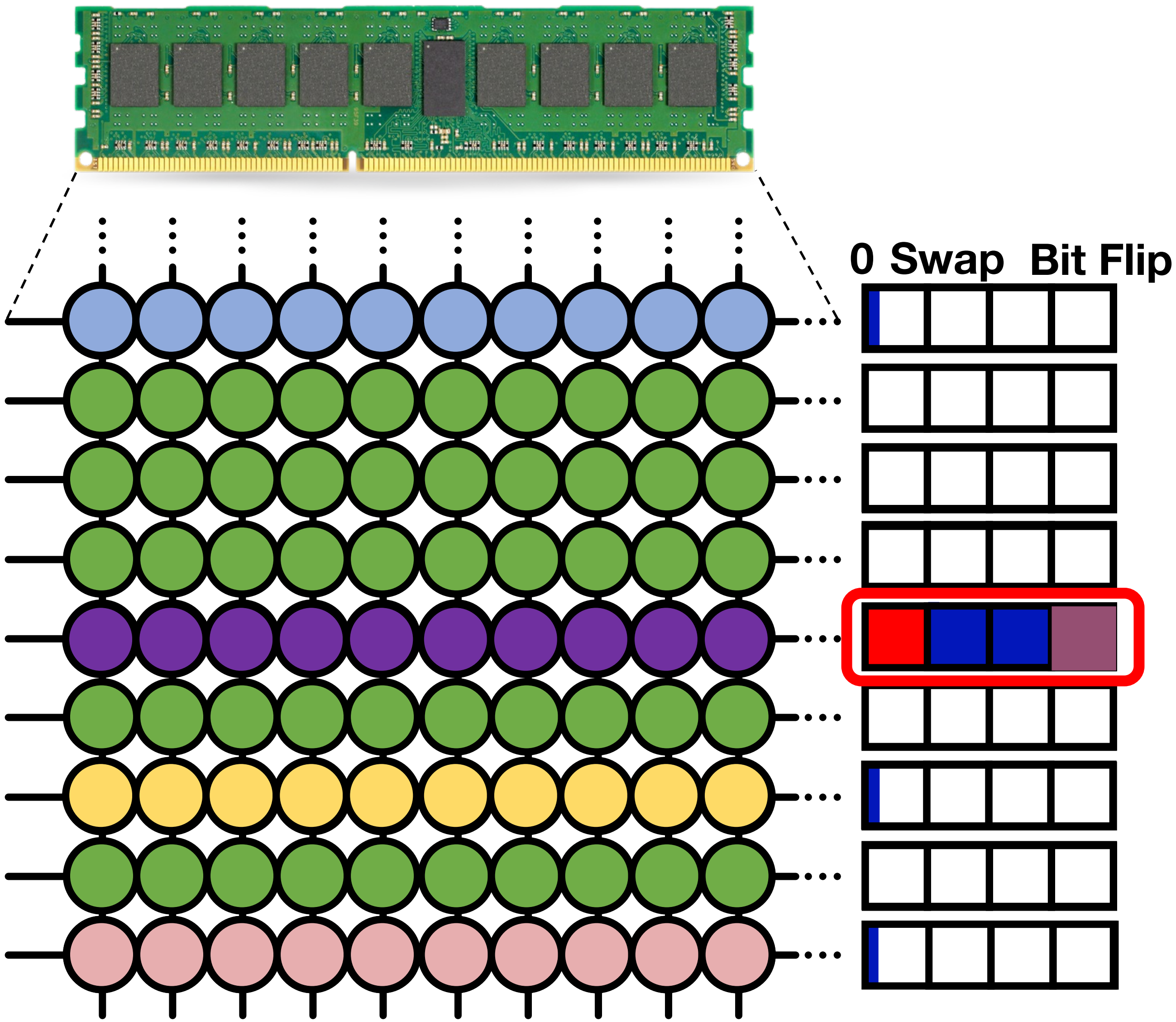
Secure Row Swap



Randomized Row-Swap (RRS)

Secure Row-Swap (SRS)

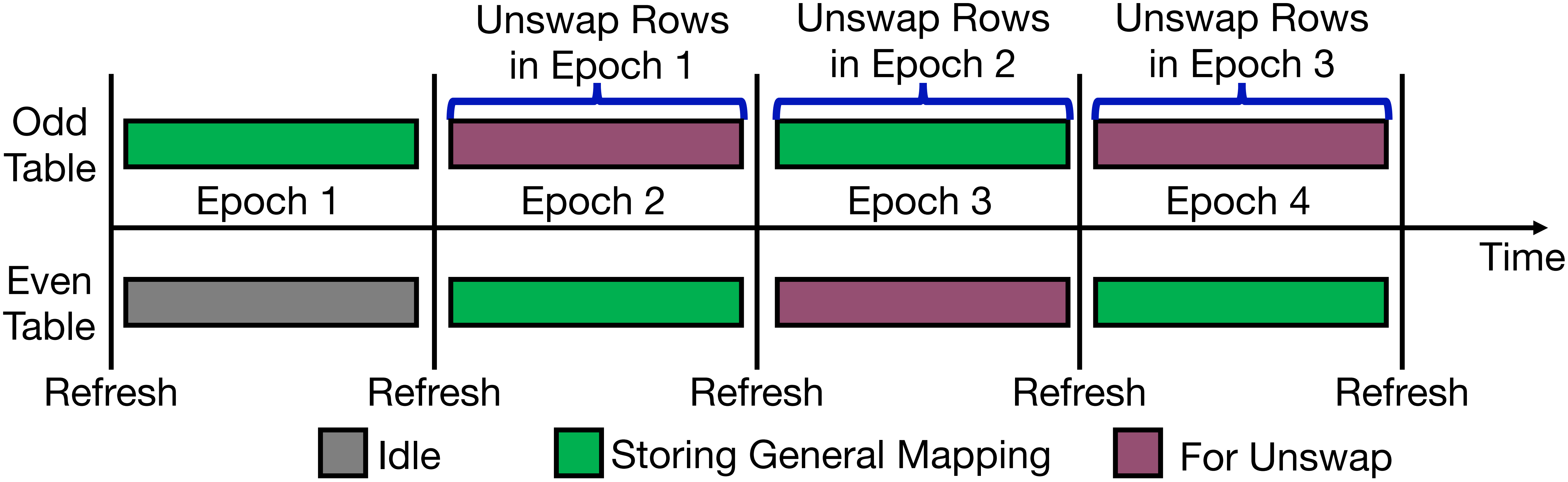
Latency Spike



Performance Degradation

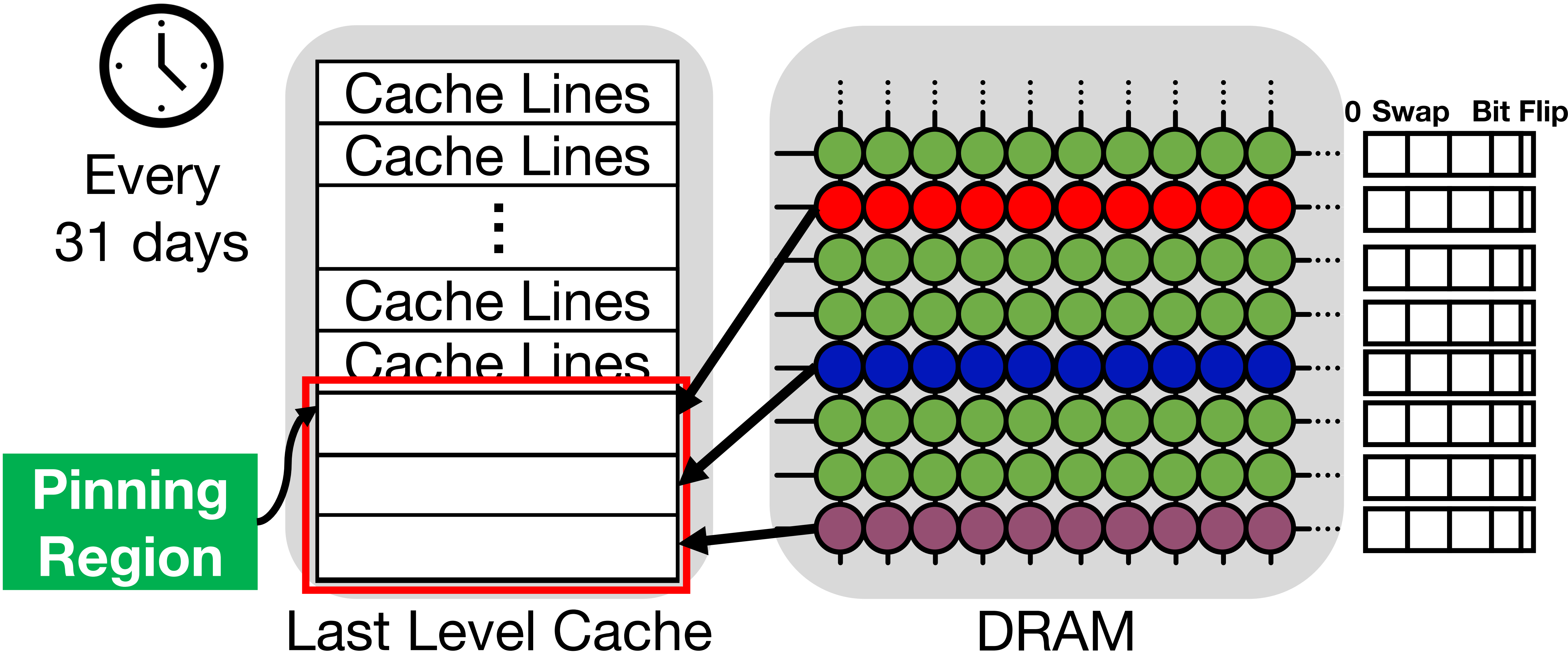
Secure Row-Swap (SRS)

Key Idea: Delay unswaps using two separate tables



✓ **No Performance Overhead Due to Unswaps**

Scalable and Secure Row-Swap



Every 31 days

Pinning Region

Cache Lines

Cache Lines

⋮

Cache Lines

Cache Lines

Last Level Cache

0 Swap

Bit Flip

DRAM

Reduces the Swap Threshold from TRH/6 to TRH/3

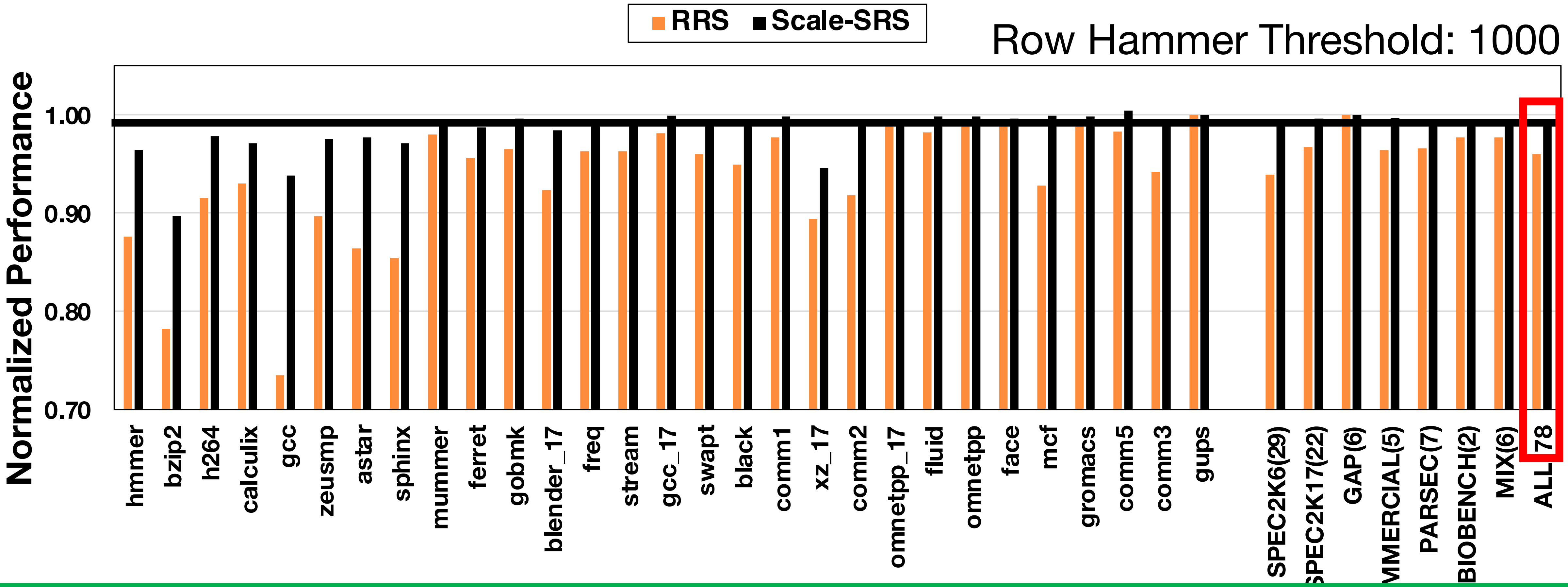
Scale-SRS: SRAM Overhead

Overhead Per Bank for TRH = 1K

Row Hammer Threshold	RRS	Scale-SRS
4800	36 KB	18.7 KB
2400	131 KB	44.4 KB
1200	251 KB	76.9 KB

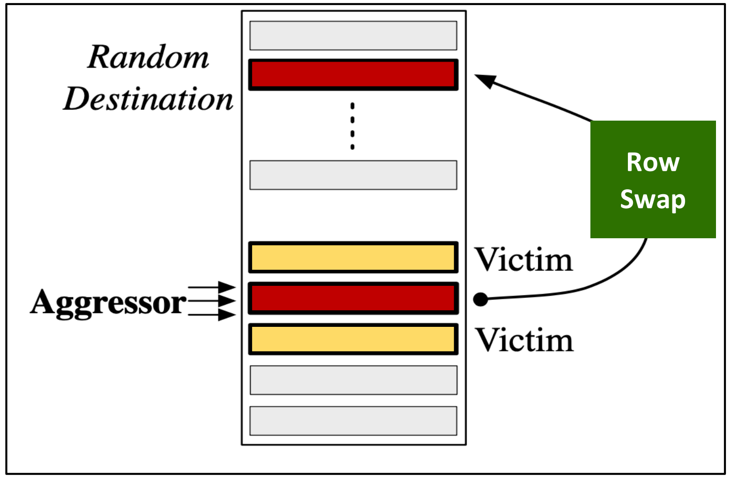
✓ **3.3X lower SRAM Overhead**

Scale-SRS: Performance



Less than 1% performance overhead

Takeways from Secure Row Swap



Enables a Secure Implementation of a Aggressor-Focused Mitigation

Scalable Solution at TRH of 1K (Less than 1% Slowdown, 70KB SRAM/bank)

Agenda

Introduction

New Mitigative Actions for Rowhammer

Randomized Row-Swaps [ASPLOS 2022, HPCA 2023]

Secure In-DRAM Tracking

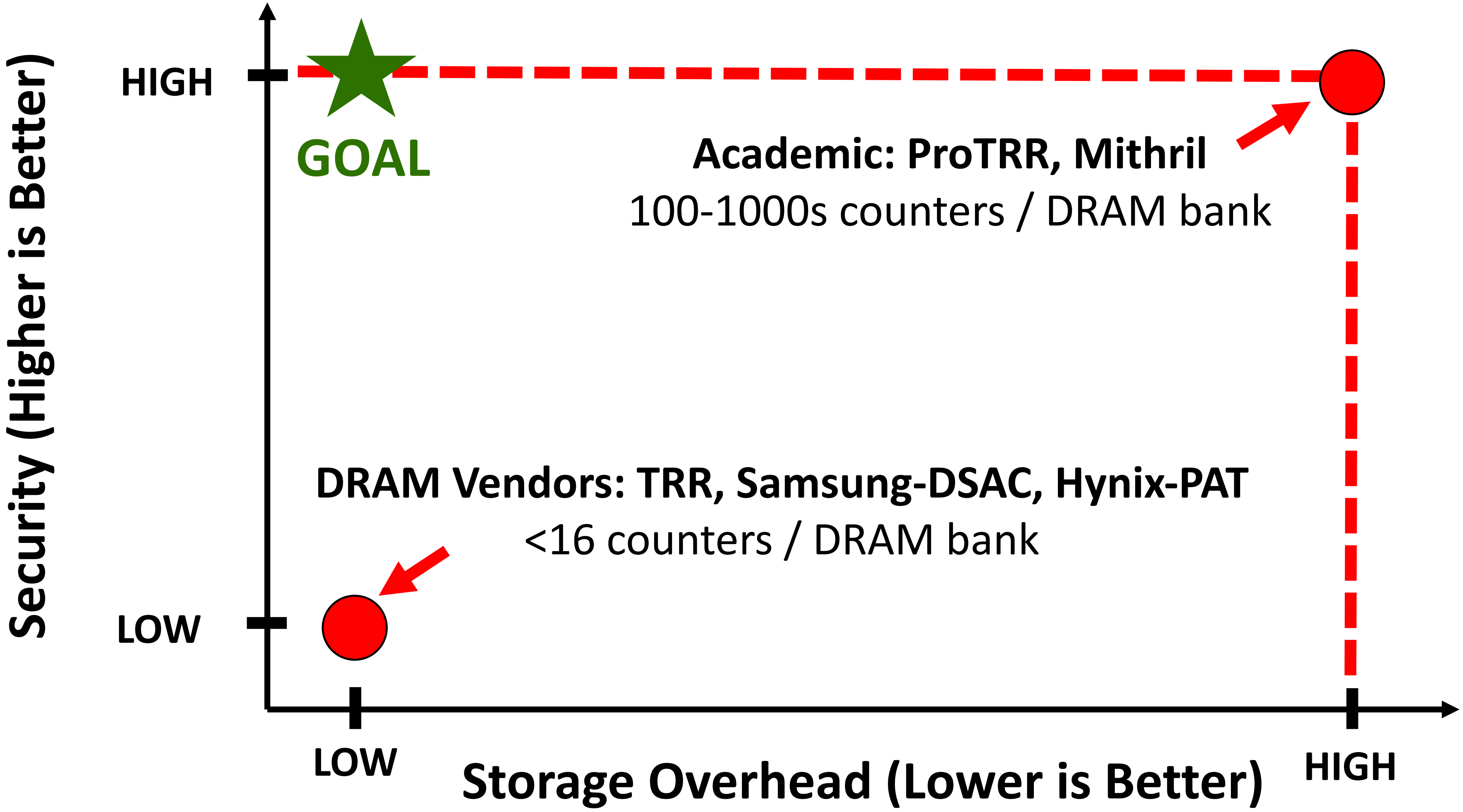
PrIDE: Probabilistic In-DRAM Tracker [ISCA 2024]

Defense in Depth Solutions

PT-Guard [DSN 2023]

Conclusion

Problem: Commercial In-DRAM Trackers Insecure



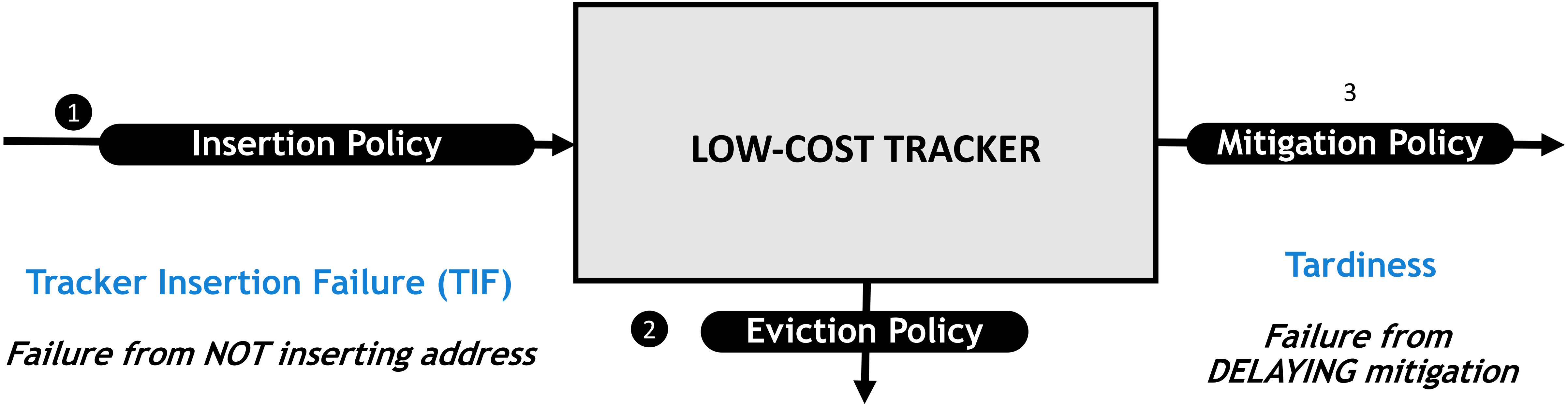
PrIDE: Achieving Secure RowHammer Mitigation Using Low Cost In-DRAM Trackers

ISCA 2024

Aamer Jaleel (NVIDIA), Gururaj Saileshwar (Toronto),
Steve Keckler (NVIDIA), Moinuddin Qureshi (GT)

Why Do Existing Low Cost In-DRAM Trackers Fail?

Taxonomy of Tracker Management Policies and Failure Modes



1 Insertion Policy

LOW-COST TRACKER

3 Mitigation Policy

2 Eviction Policy

Tracker Insertion Failure (TIF)

Failure from NOT inserting address

Tracker Retention Failure (TRF)

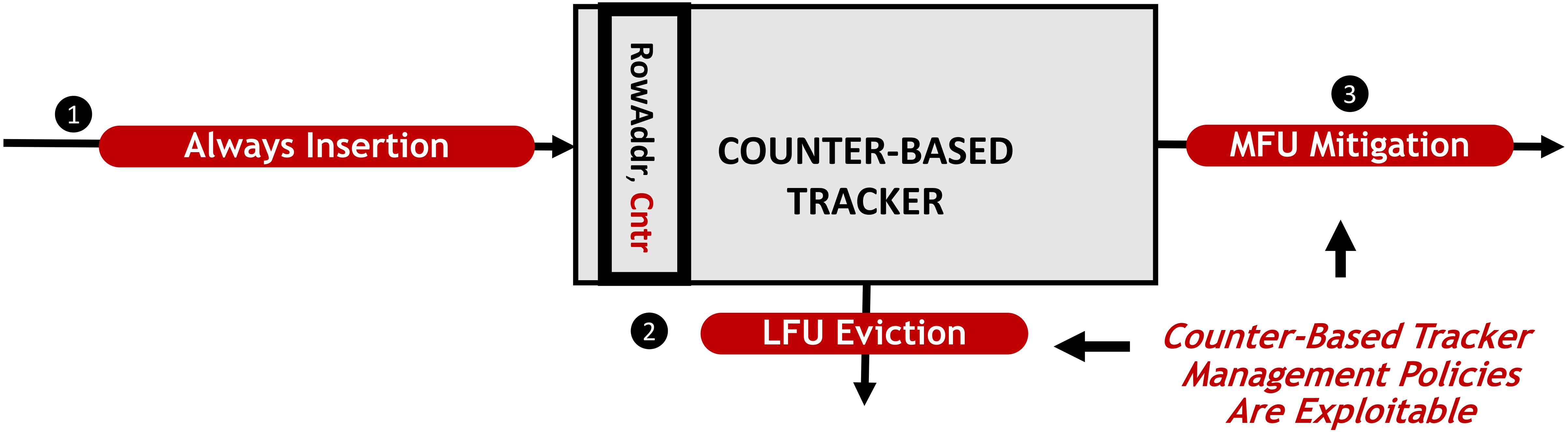
Failure from DISCARDING address

Tardiness

Failure from
DELAYING mitigation

Why Do Existing Low Cost In-DRAM Trackers Fail?

Current Trackers Use Counters to Track Frequently Activated Rows

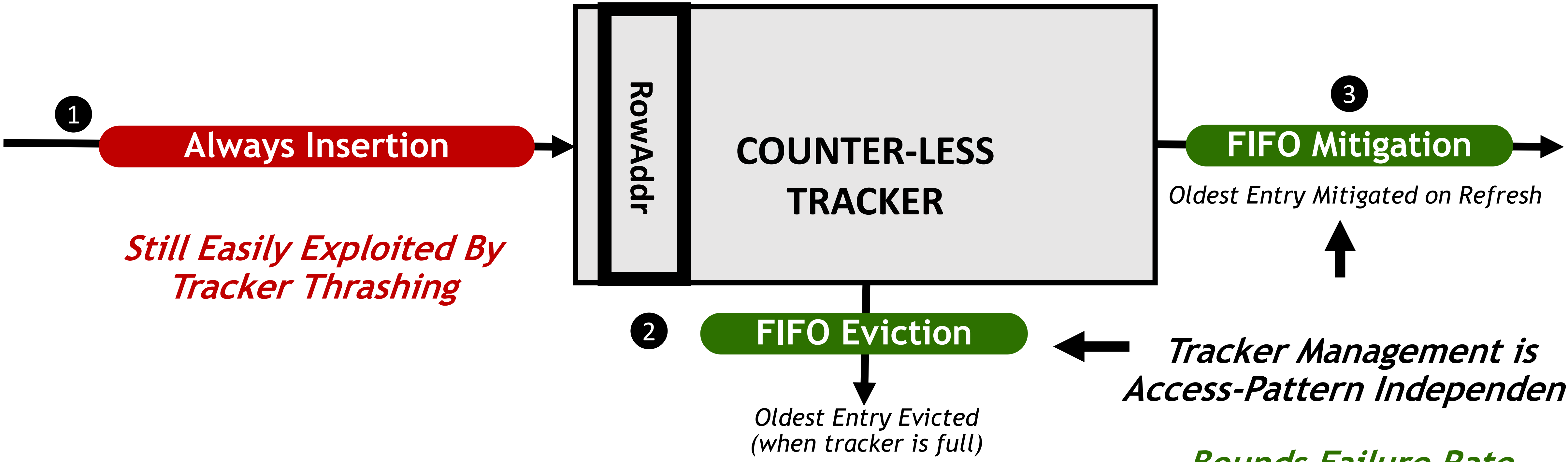


Counter-Based Tracker Management Policies Are Exploitable

Existing Tracker Policies are Access Pattern Dependent

Carefully Crafted Access Patterns (e.g. TRRespass) Induce Tracker Retention Failures!

Insight: Secure In-DRAM Tracker Requires Access-Pattern Independence

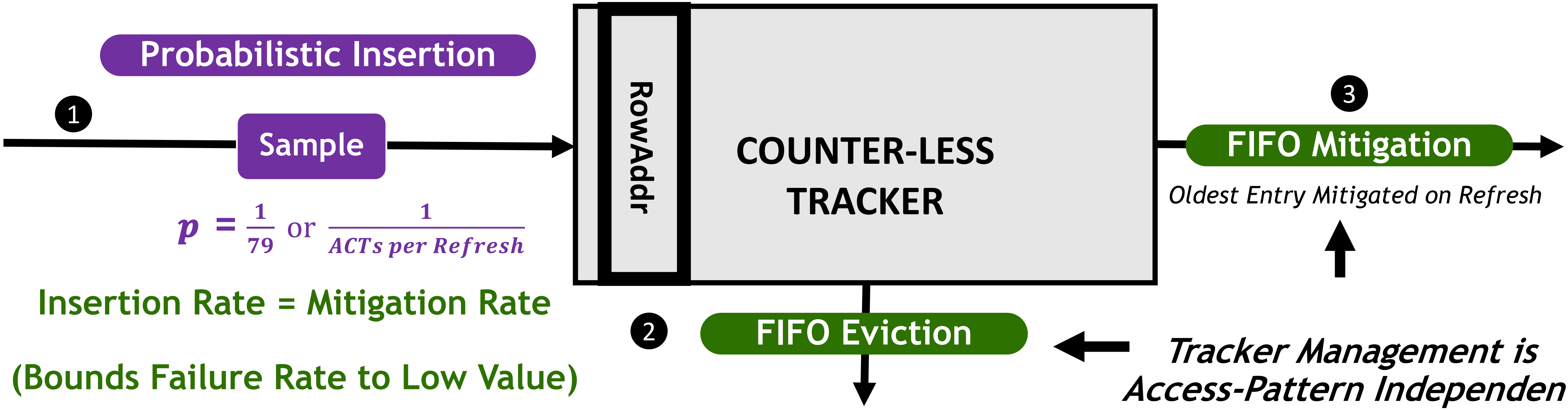


Still Easily Exploited By Tracker Thrashing

Tracker Management is Access-Pattern Independent

Bounds Failure Rate

PrIDE = Probabilistic Insertion + Access-Pattern Independent Tracker Management



$$p = \frac{1}{79} \text{ or } \frac{1}{ACTs\ per\ Refresh}$$

Insertion Rate = Mitigation Rate

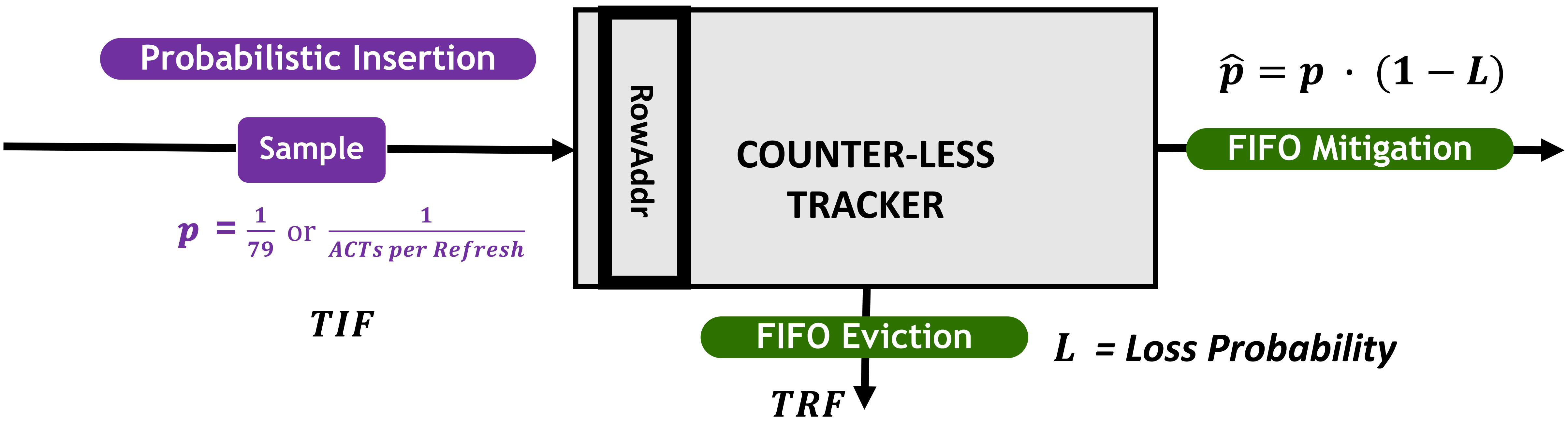
(Bounds Failure Rate to Low Value)

Oldest Entry Mitigated on Refresh

Tracker Management is Access-Pattern Independent

Bounds Failure Rate

PrIDE = Probabilistic Insertion + Access-Pattern Independent Tracker Management



$$p = \frac{1}{79} \text{ or } \frac{1}{ACTs\ per\ Refresh}$$

$$\hat{p} = p \cdot (1 - L)$$

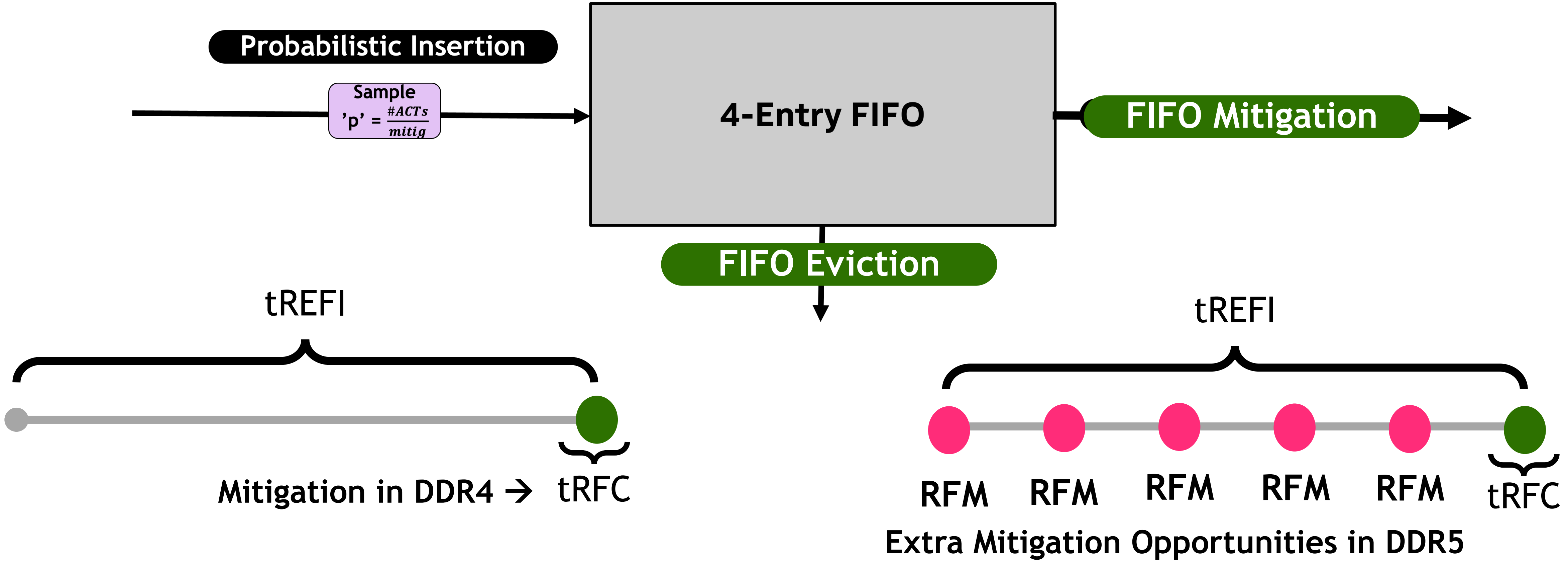
L = *Loss Probability*

$$MTTF = TIF + TRF = (1 - p \cdot (1 - L))^{TRH}$$

Mathematical Proof & Analysis in the Paper

For a MTTF of 10,000 years, PrIDE can support TRH = 1575 with 16-entry FIFO

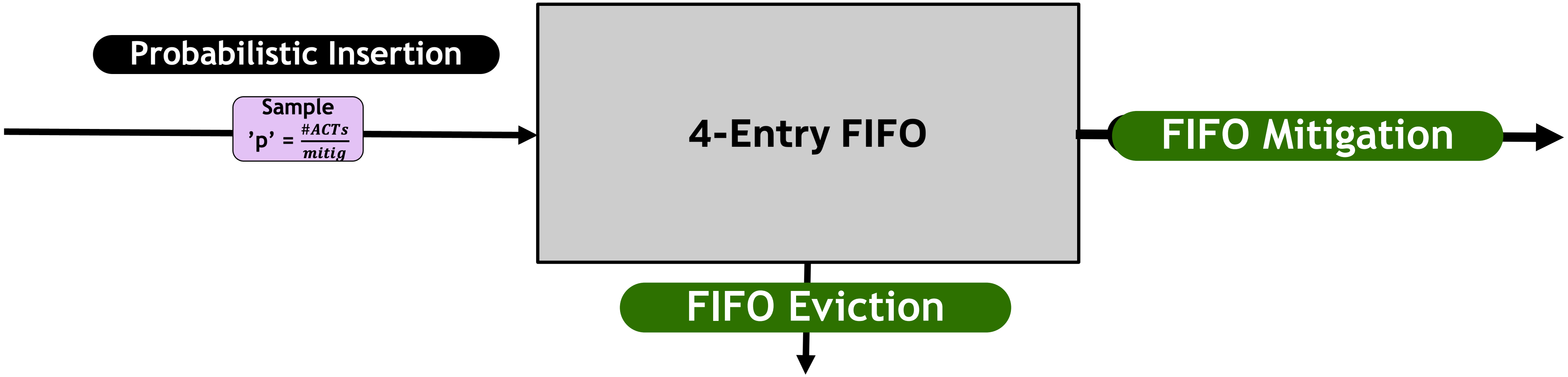
Benefits of PrIDE – Secure & Low Cost In-DRAM Tracker



PrIDE (DDR4):
Performance Overhead: 0%
TRH = 1900

PrIDE + RFM16 (DDR5):
Performance Overhead: 1.6%
TRH = 400

Takeways from PrIDE



FIRST Low-Cost and Secure In-DRAM Defense for Future DRAM

Scalable to TRH of 400 at Negligible Cost (1% Slowdown, 16 Bytes SRAM/bank)

Agenda

Introduction

New Mitigative Actions for Rowhammer

Randomized Row-Swaps [ASPLOS 2022, HPCA 2023]

Secure In-DRAM Tracking Solutions

PrIDE: Probabilistic In-DRAM Tracker [ISCA 2024]

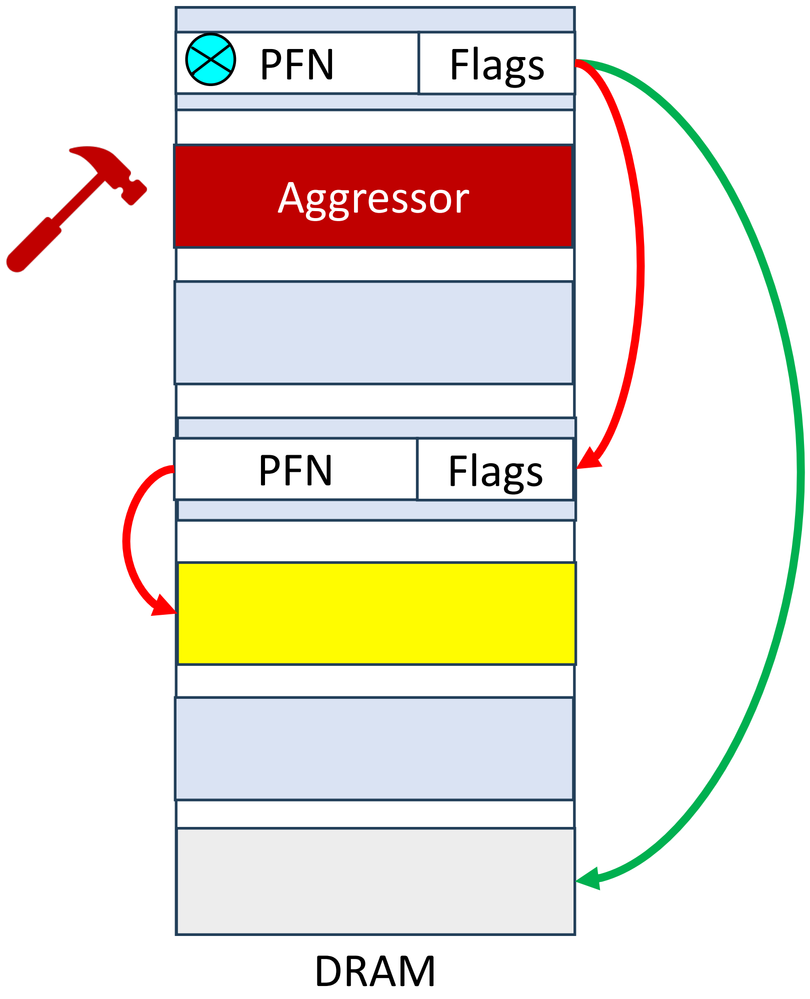
Defense in Depth Solutions

PT-Guard [DSN 2023]

Conclusion

Be Paranoid: Defenses can be Broken

Privilege Escalation Exploit with Rowhammer



🟦 Attacker’s page table

⬜ Attacker accessible page

🟨 Attacker inaccessible page

Bit-Flips in page tables enables privilege escalation, breaking system security

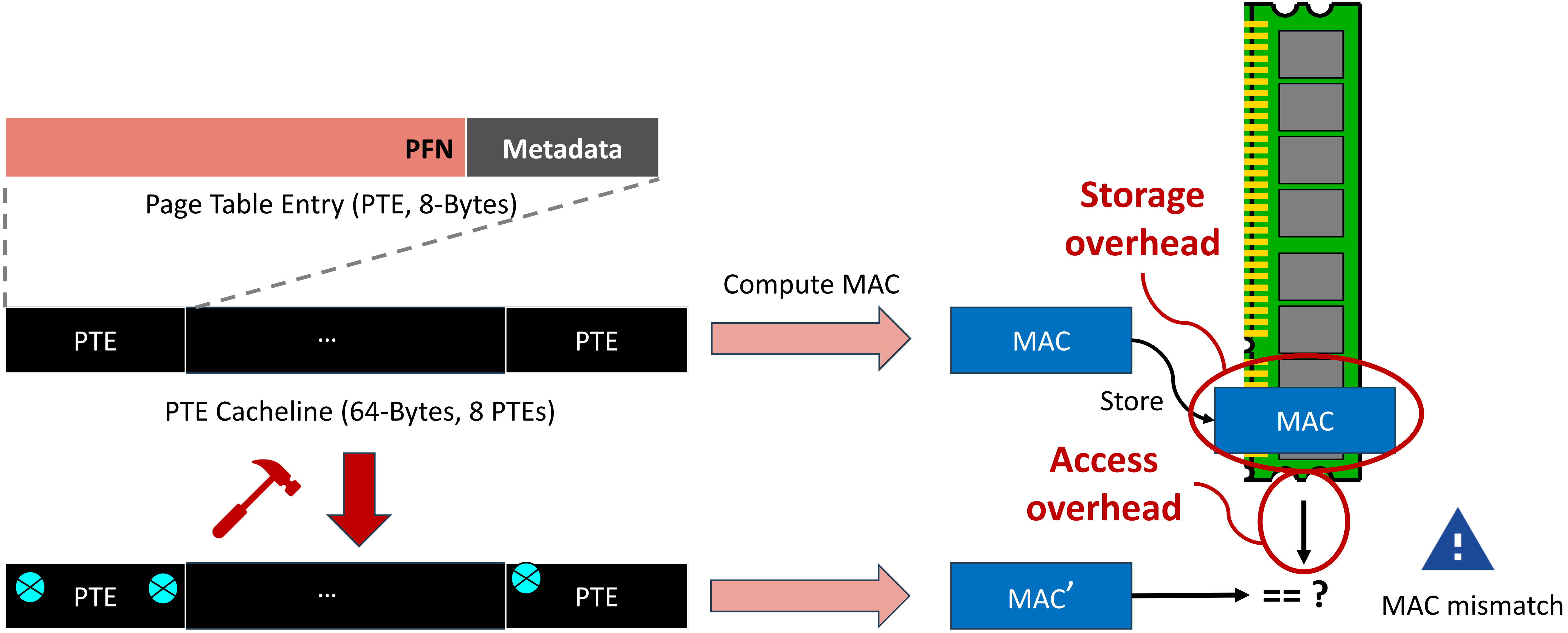
PT-Guard: Integrity-Protected Page Tables to Defend Against Breakthrough Rowhammer Attacks

DSN 2023, Spain

Anish Saxena, Gururaj Saileshwar, Jonas Juffinger,
Andreas Kogler, Daniel Gruss , Moinuddin Qureshi

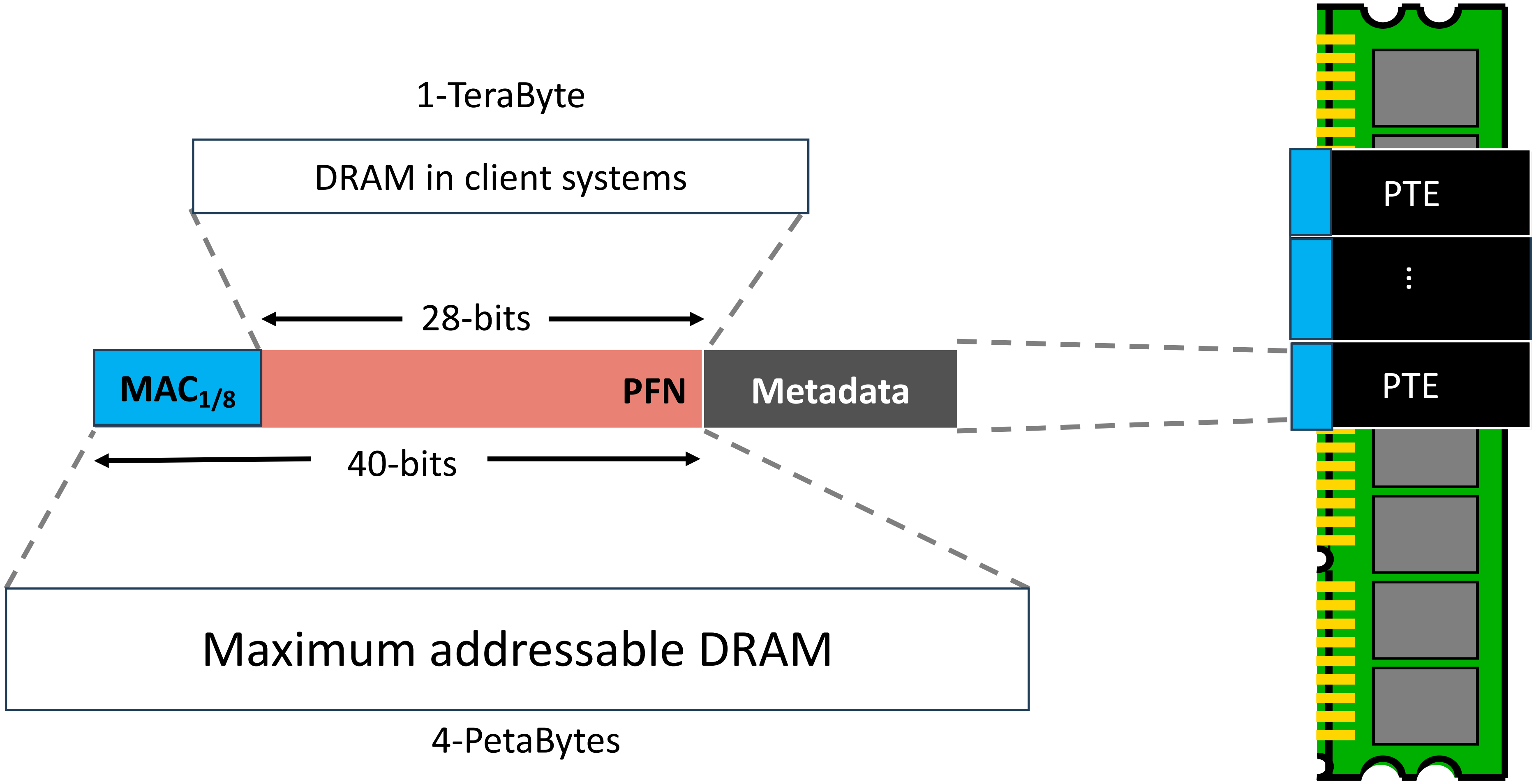
Georgia Tech NVIDIA TU Graz

Mac-Based Integrity Protection



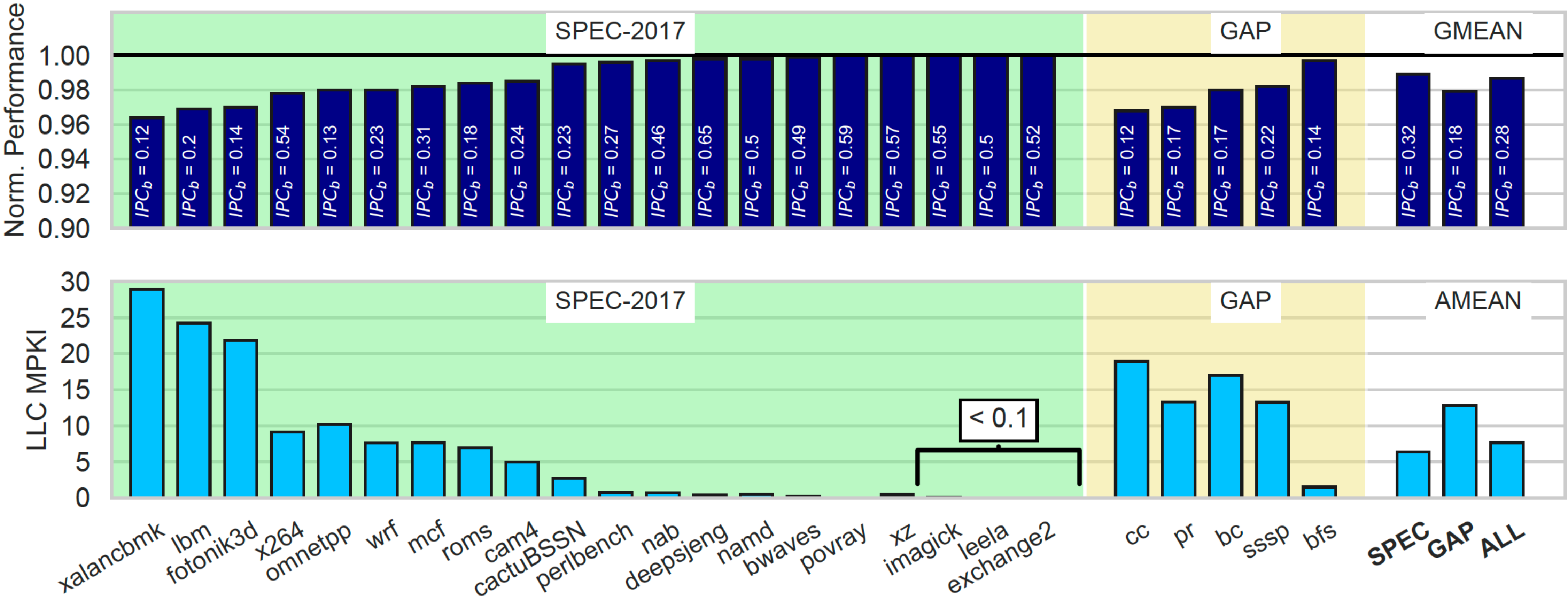
MAC provides cryptographic integrity protection but has high overheads

Embedding MAC within the PTE cacheline



PT-Guard embeds a 96-bit MAC within the PTE line, obviating storage and access overheads

Evaluation Results



PT-Guard provides integrity for page tables at 1.3% slowdown.

PT-Guard also has best-effort correction (corrects about 90% of errors at 0.5% bit error rate)

Agenda

Introduction

New Mitigative Actions for Rowhammer

Randomized Row-Swaps [ASPLOS 2022, HPCA 2023]

Secure In-DRAM Tracking

PrIDE: Probabilistic In-DRAM Tracker [ISCA 2024]

Defense in Depth Solutions

PT-Guard [DSN 2023]

Conclusion

Conclusions

Rowhammer Vulnerability is Becoming Worse!

- New attack patterns likely to emerge as attacker capability increases.

Defenses Need to be Practical & Resilient to Old & New Attacks

- RRS, SRS → New Mitigative Actions focused on Aggressors
- PrIDE → First Secure and Low-Cost In-DRAM Defense
- PTGuard → Defense in Depth

Looking Forward: Long Way to Go!

- Explore Threat Landscape on Emerging DRAM (DDR5, HBM, GDDR)
- Make Critical SW Applications (e.g., ML Models) Resilient to Rowhammer
- Address Vulnerability at Low-Cost in Future DRAM (sub-100 thresholds)

Thank you! Questions?

UNIVERSITY OF TORONTO

DEFY GRAVITY