Rowhammer Attacks

Mengjia Yan
mengjia@csail.mit.edu

Based on slides by Onur Mutlu SEHAS Keynote @ HiPEAC

The “CIA Triad”

- Confidentiality
  - Prevent sensitive data from unauthorized access
  - *E.g. Side-channel attacks*
- Availability
  - Ensure that information is consistently and readily accessible to authorized parties
  - *E.g. DoS attacks*
- Integrity
  - Maintain the consistency, accuracy, and trustworthiness of data
  - E.g. Fault Injection, **Rowhammer**

The Story of RowHammer

- One can predictably induce bit flips in commodity DRAM chips
- An example of how a simple hardware failure mechanism can create a widespread system security vulnerability

WIRED

Forget Software—Now Hackers Are Exploiting Physics

BUSINESS | CULTURE | DESIGN | GEAR | SCIENCE

SHARE

SHARE 18276

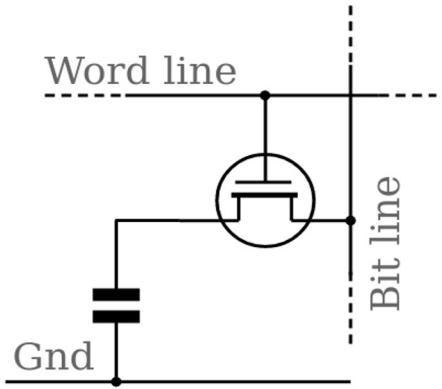
TWEET

ANDY GREENBERG SECURITY 08.31.16 7:00 AM

FORGET SOFTWARE—NOW HACKERS ARE EXPLOITING PHYSICS

DRAM Basics

- Each bit in DRAM is stored in a “cell” using a *capacitor*
- *Read is destructive*
- DRAM cells lose their state over time (hence ***Dynamic*** RAM)
  - Data stored in DRAM cells needs to be “refreshed” at a regular interval



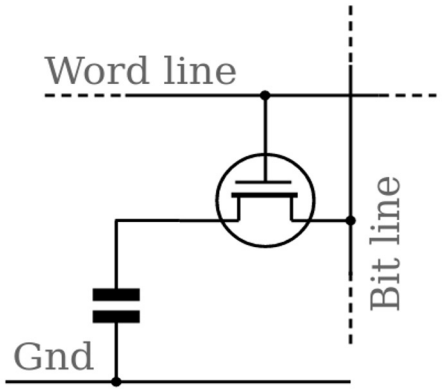
Why we widely use DRAM?

- Speed
- Density
- Cost

https://www.electronics-notes.com/articles/electronic_components/semiconductor-ic-memory/dynamic-ram-how-does-dram-work-operation.php

DRAM Basics

- Each bit in DRAM is stored in a “cell” using a *capacitor*
- *Read is destructive*
- DRAM cells lose their state over time (hence ***Dynamic*** RAM)
  - Data stored in DRAM cells needs to be “refreshed” at a regular interval

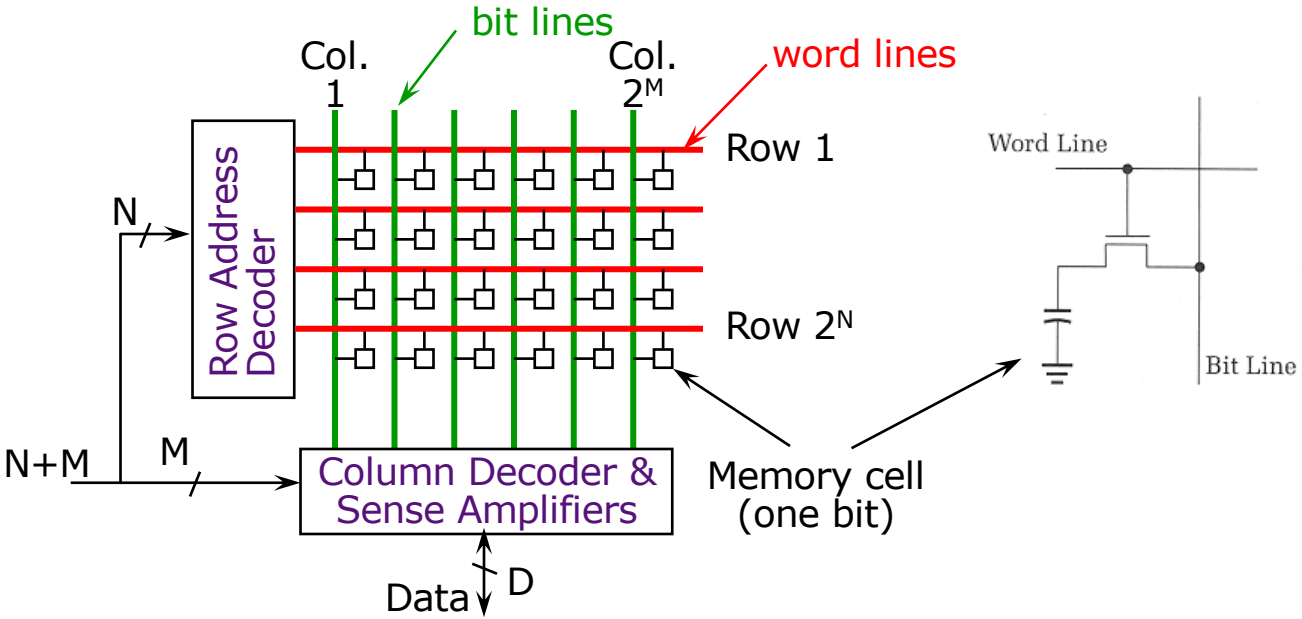


Why we widely use DRAM?

- Speed (2-10x slower than SRAM)
- Density (20x denser than SRAM)
- Cost (~100x cheaper per MB)

https://www.electronics-notes.com/articles/electronic_components/semiconductor-ic-memory/dynamic-ram-how-does-dram-work-operation.php

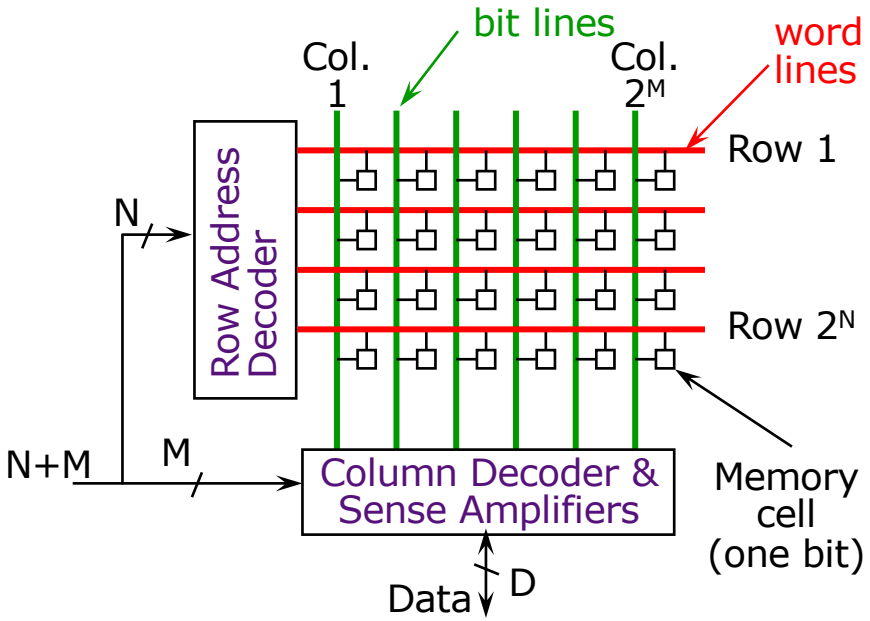
DRAM Architecture



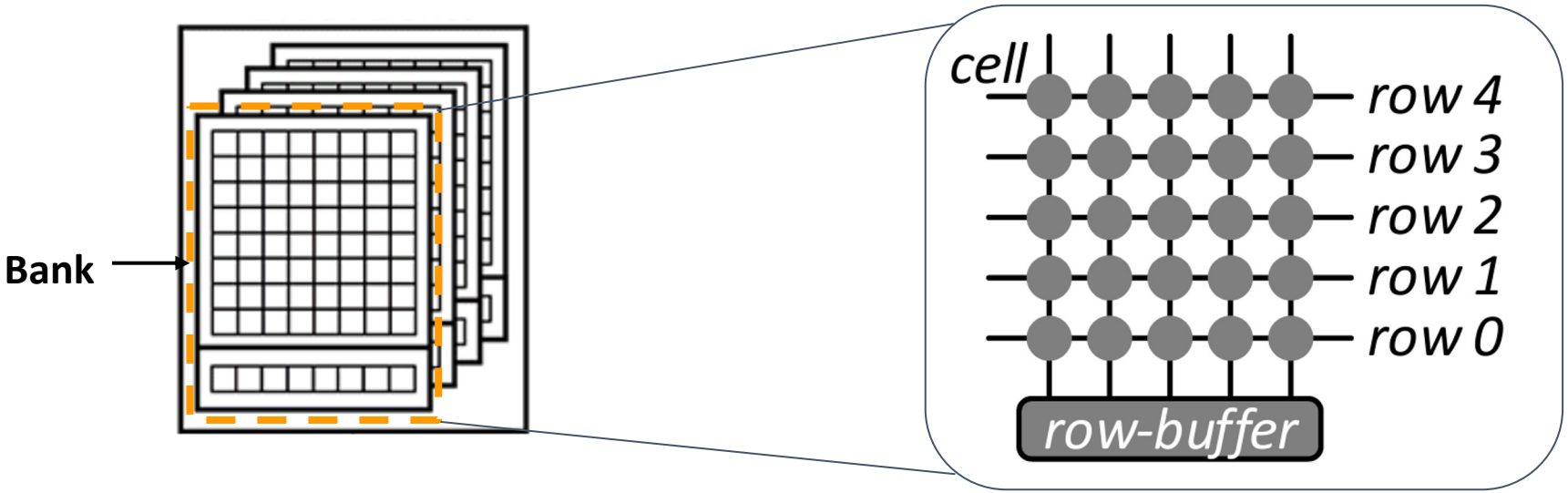
- Bits stored in 2-dimensional arrays on chip
- Question: why read the entire row?

DRAM Refresh

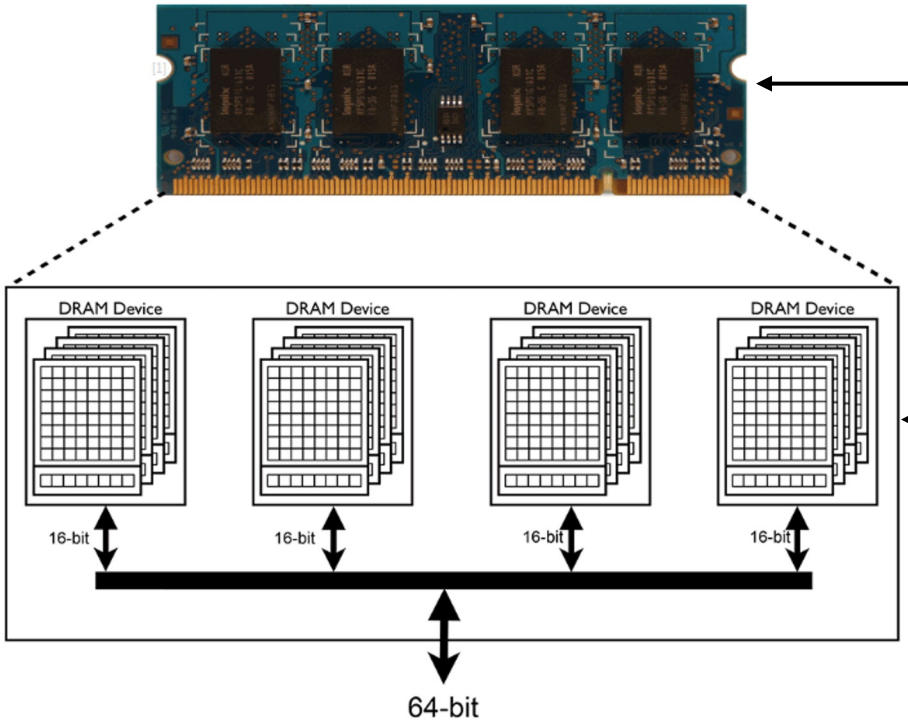
- How to do refresh?
- Performance penalty of refresh
  - In an 8Gb memory, upwards of 10% of time is spent in refresh!
- The common refresh interval: **64ms**



DRAM Organization



DRAM Organization



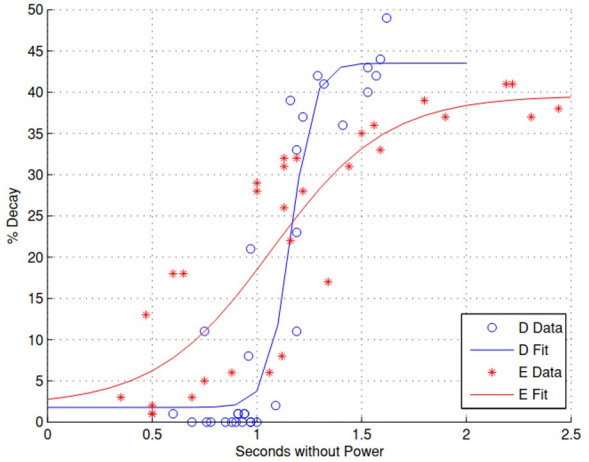
DIMM
Dual Inline Memory Module

Rank

https://apt.cs.manchester.ac.uk/projects/ARMOR/RowHammer/index.html

Aside: Cold Boot Attacks

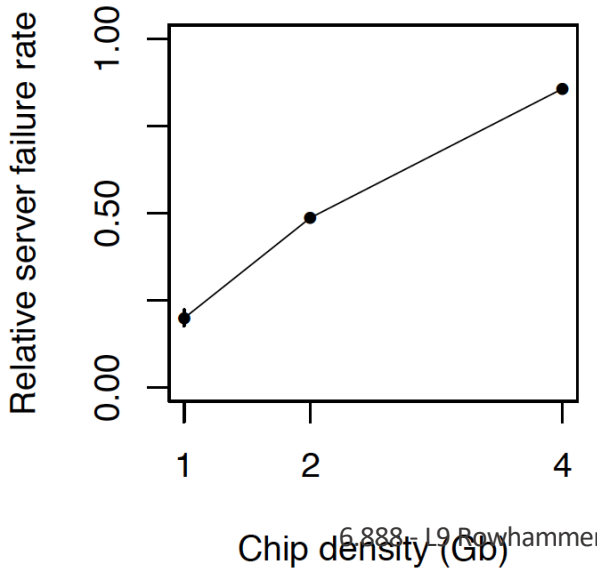
	Seconds w/o power	Error % at operating temp.	Error % at −50 °C
SDRAM (1999)	60	41	(no errors)
	300	50	0.000095
DDR (2001)	360	50	(no errors)
	600	50	0.000036
DDR (2003)	120	41	0.00105
	360	42	0.00144
DDR2 (2007)	40	50	0.025
	80	50	0.18



Halderman et al.; Lest We Remember: Cold Boot Attacks on Encryption Keys; USENIX Security'08

As Memory Scales, It Becomes Unreliable

- Data from all of Facebook's servers worldwide
- Meza+, "Revisiting Memory Errors in Large-Scale Production Data Centers," DSN'15.



Error Correcting Codes (ECC)

- **Basic Idea:** Store extra *redundant* bits to be used in case of a flip!
- **Naive Implementation:** Store multiple copies and compare
  - Expensive!
- **Actual Implementation:** Hamming codes!

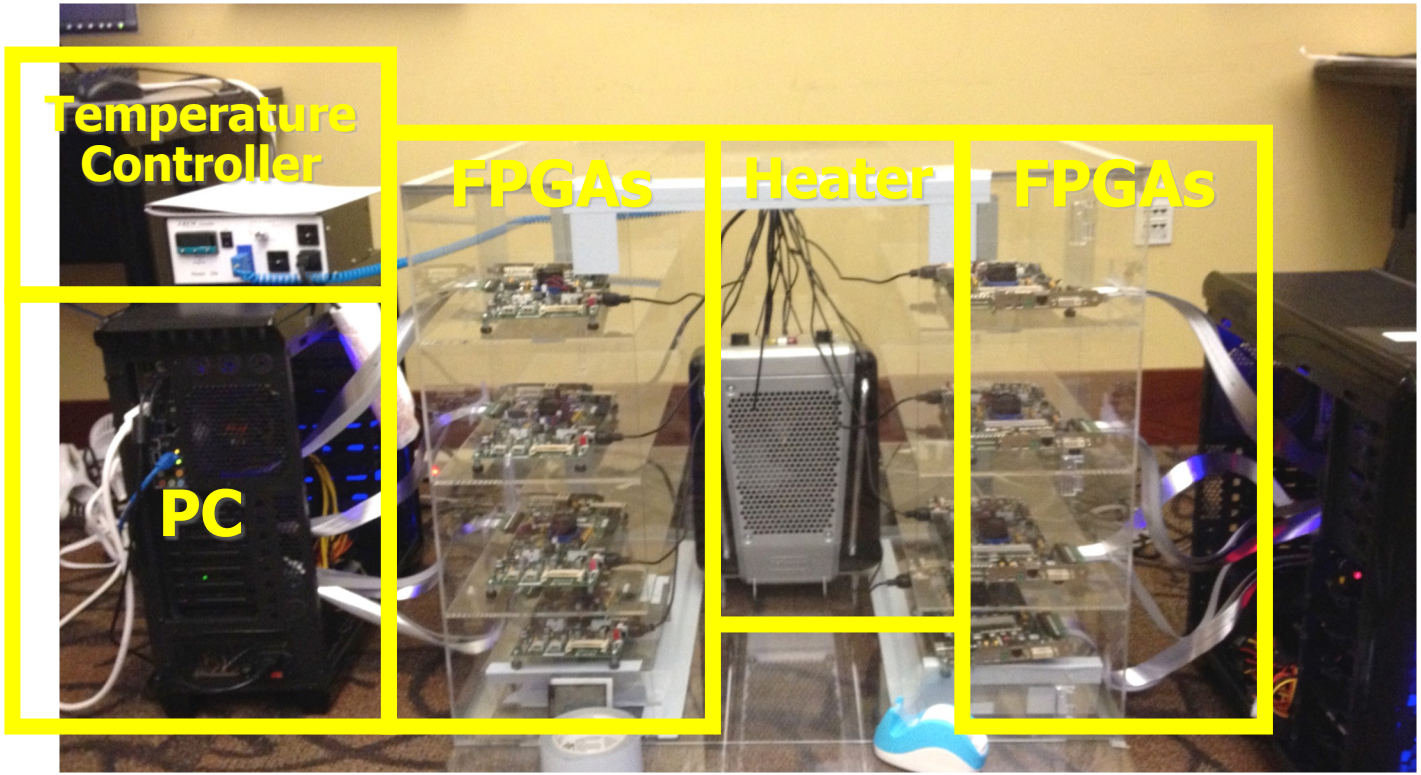
Hamming codes allow for *single-error* correction, double error detection (aka **SECDED**)

How about more than 2-bit flips?

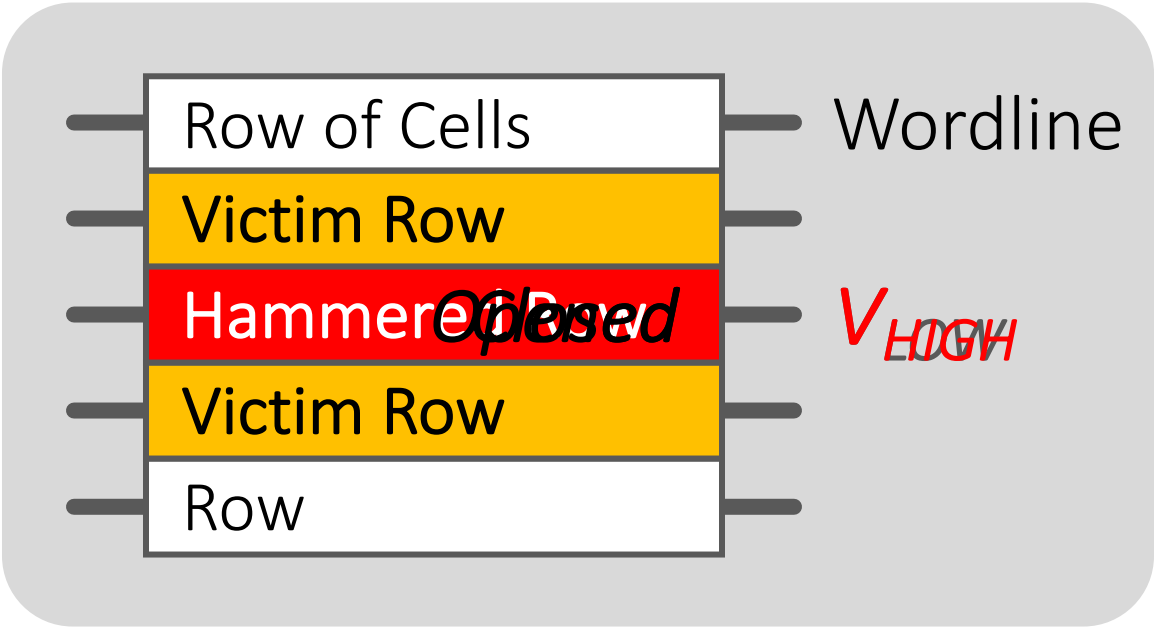
Reliability + Security Implications

Robust Physical-World Attacks on Deep Learning Visual Classification - Eykholt et al.

Infrastructures to Understand Such Issues



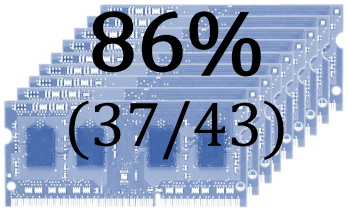
Kim et al; Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors; ISCA'14



Observation: Repeatedly accessing a row enough times between refreshes can cause disturbance errors in nearby rows

Most DRAM Modules Are Vulnerable

A company

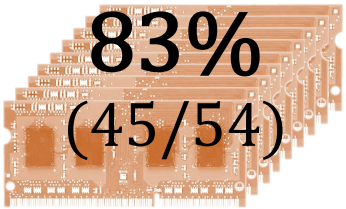


Up to

1.0×10^7

errors

B company

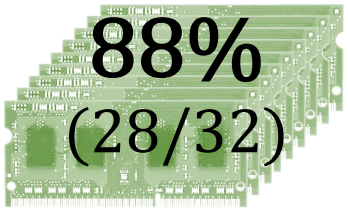


Up to

2.7×10^6

errors

C company



Up to

3.3×10^5

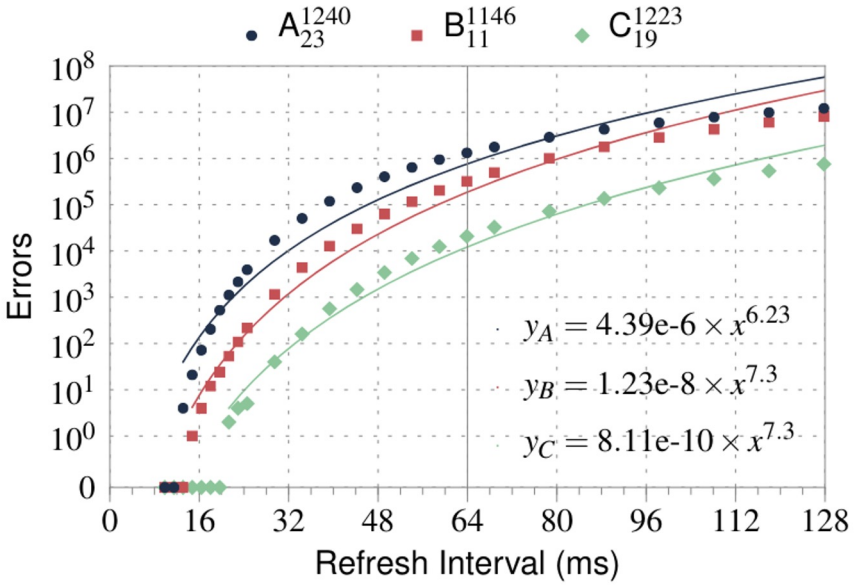
errors

Why Is This Happening?

- DRAM cells are too close to each other!
  - They are not electrically isolated from each other
- Access to one cell affects the value in nearby cells
  - Due to electrical interference between the cells and wires used for accessing the cells
  - Also called cell-to-cell coupling/interference
- Example: When we activate (apply high voltage) to a row, an adjacent row gets slightly activated as well
  - Vulnerable cells in that slightly-activated row lose a little bit of charge
  - If row hammer happens enough times, charge in such cells gets drained

Refresh + Hammering Interval Effects

Examining error rates for different refresh and hammering rates on DDR2 modules from 2011-2012



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

Apple's Patch for RowHammer

- https://support.apple.com/en-gb/HT204934

  Available for: OS X Mountain Lion v10.8.5, OS X Mavericks v10.9.5

  Impact: A malicious application may induce memory corruption to escalate privileges

  Description: A disturbance error, also known as Rowhammer, exists with some DDR3 RAM that could have led to memory corruption. This issue was mitigated by increasing memory refresh rates.

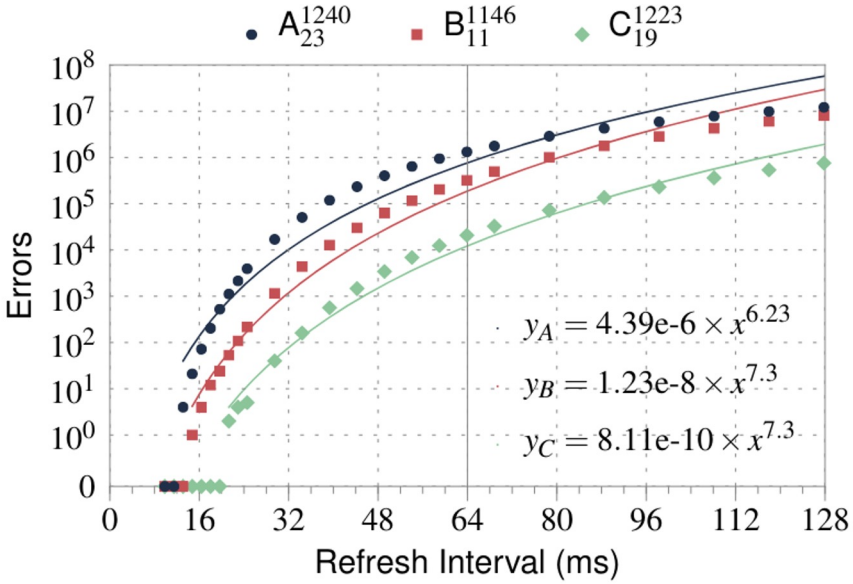
  CVE-ID

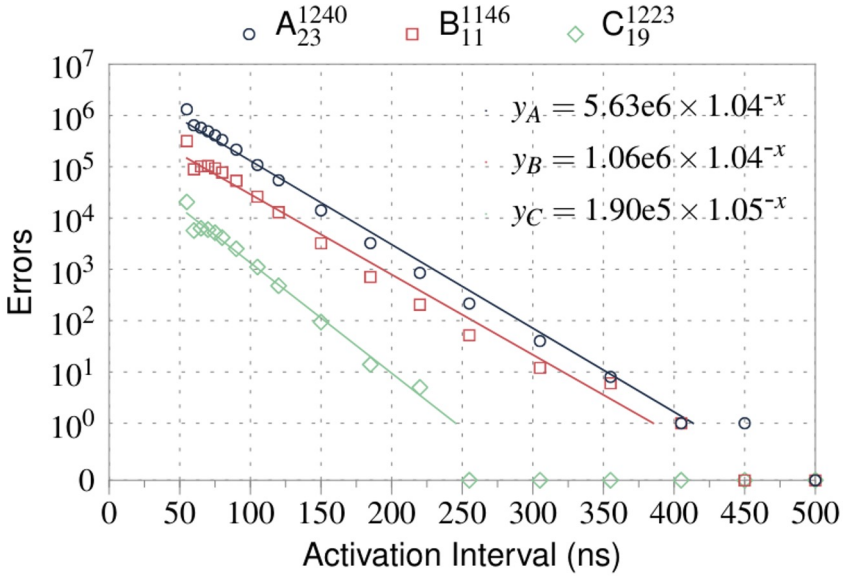
  CVE-2015-3693 : Mark Seaborn and Thomas Dullien of Google, working from original research by Yoongu Kim et al (2014)

HP, Lenovo, and many other vendors released similar patches

Refresh + Hammering Interval Effects

Examining error rates for different refresh and hammering rates on DDR2 modules from 2011-2012

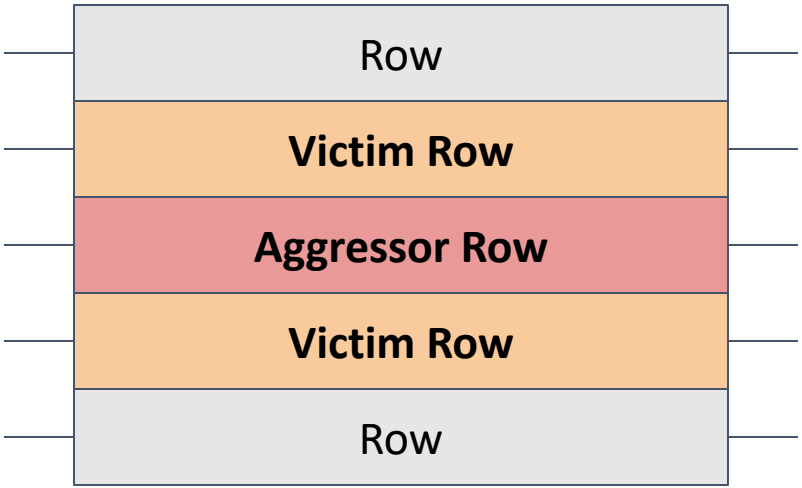




Flipping Bits in Memory Without Accessing Them: An Experimental Study of DRAM Disturbance Errors

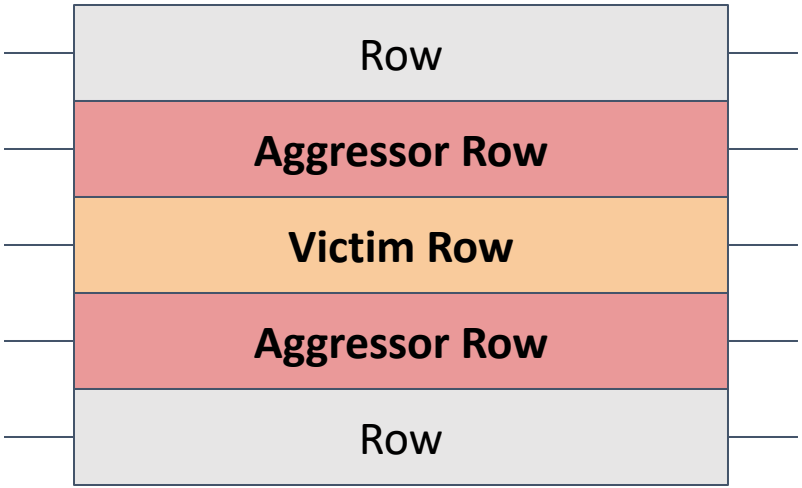
“Single-Sided” Rowhammer

- Aggressor Row = Hammered Row



Question: Can we increase the “stress” on a victim row?

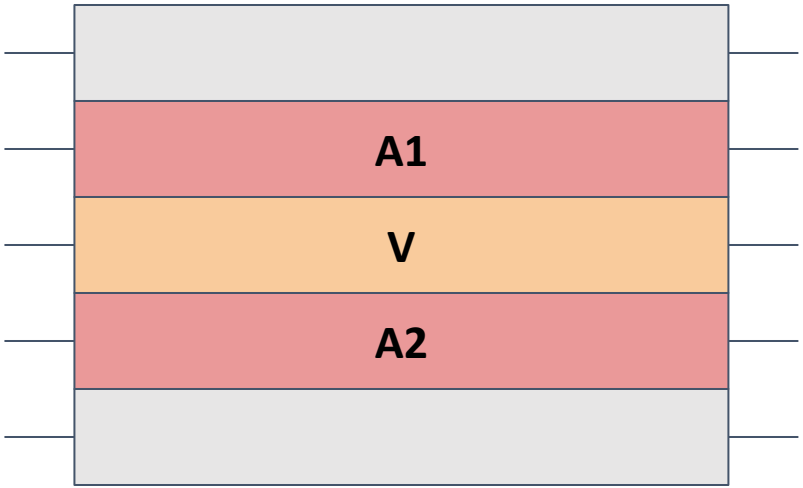
“Double-Sided” Rowhammer



Observation:

Repeatedly accessing both adjacent rows *significantly* increases the error rate of the victim row

Double-Sided Rowhammer Code Example



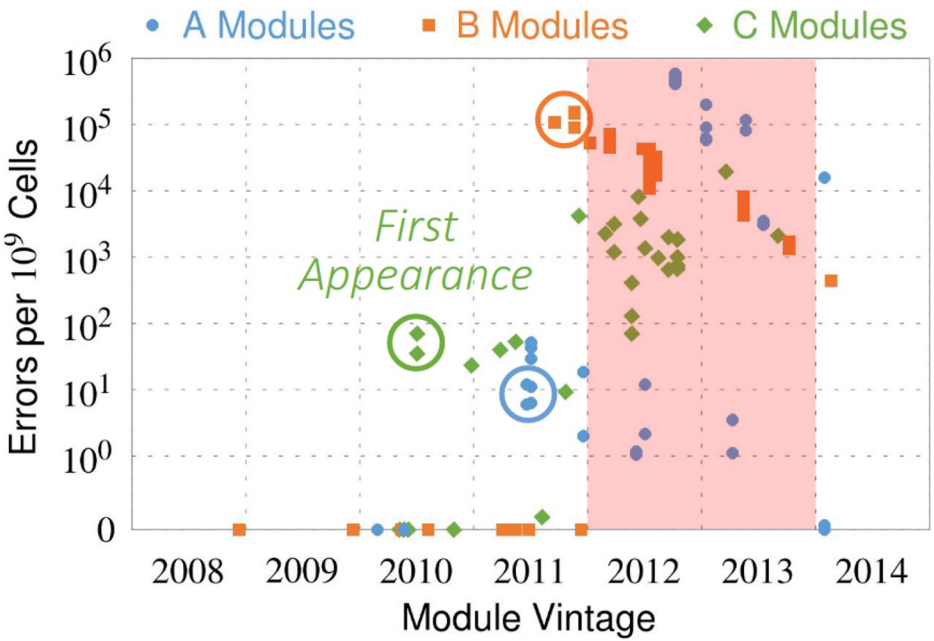

```
loop:
    mov (A1), %eax
    mov (A2), %ebx

    clflush (A1)
    clflush (A2)

    mfence
    jmp loop
```

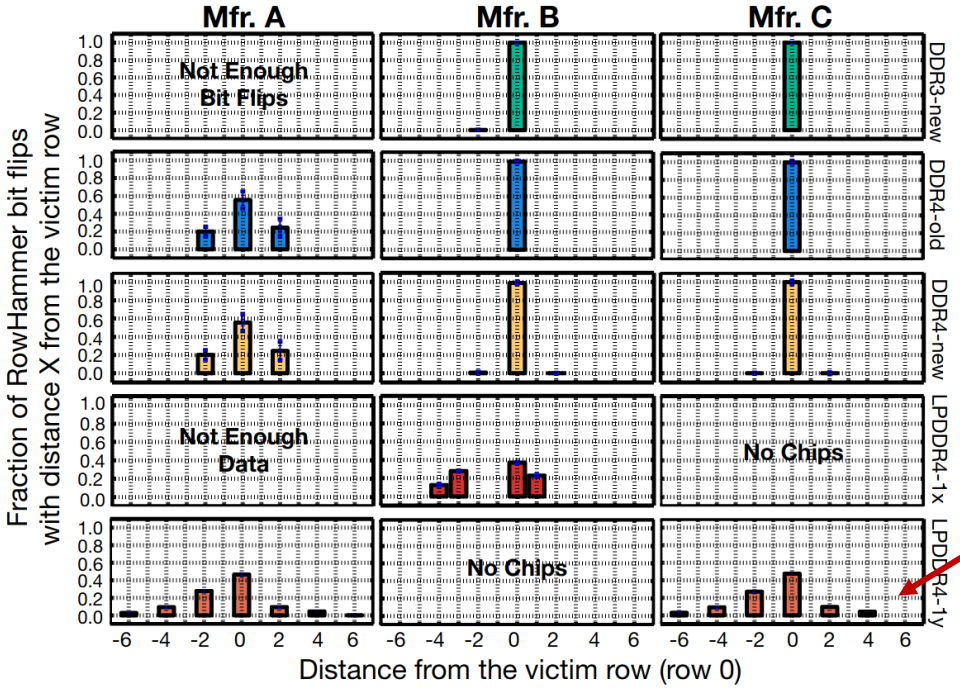
Question: Why do we need the `clflush`?

Density Trends



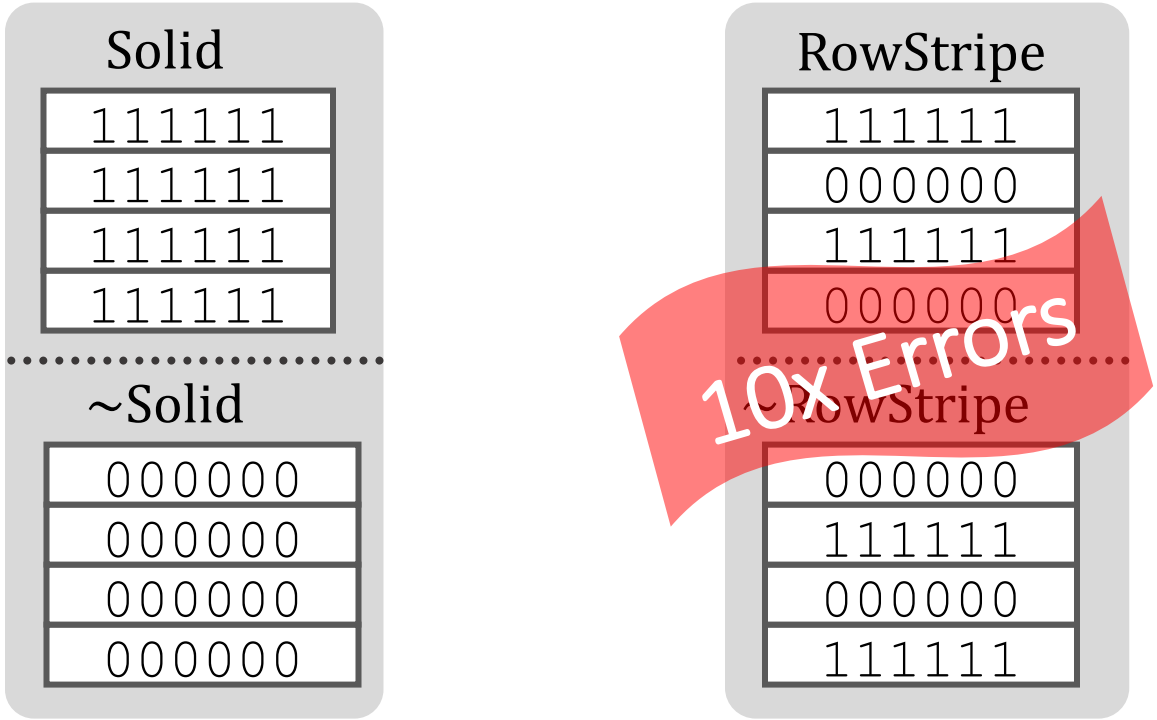
- As DRAM gets physically denser, it becomes even **more vulnerable**!
- Only a few thousand hammer iterations are required on modern DRAM to cause a bit-flip

Density Trends



Denser DRAM also can result in flips in rows which are not *directly* adjacent to the attacker

Several Findings



Errors affected by data stored in other cells

Several Other Findings

- Errors are repeatable
  - Across ten iterations of testing, >70% of victim cells had errors in every iteration
- As many as 4 errors per cache-line
  - Simple ECC (e.g., SECDED) cannot prevent all errors
- Number of cells & rows affected by aggressor
  - Victims cells per aggressor: ≤ 110
  - Victims rows per aggressor: ≤ 9

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

Abstract. *Memory isolation is a key property of a reliable and secure computing system — an access to one memory address should not have unintended side effects on data stored in other addresses. However, as DRAM process technology*

(Kim et al., ISCA 2014)

Project Zero

(Seaborn, 2015)

News and updates from the Project Zero team at Google

Monday, March 9, 2015

Exploiting the DRAM rowhammer bug to gain kernel privileges

Native Client (NaCl) Sandbox Escape

- NaCl is a sandbox for running native code (C/C++)
- Runs a “safe” subset of x86, statically verifying an executable
- Use bit flips to make an instruction sequence unsafe!

Example “Safe” Code:

```
andl $~31, %eax     // Truncate address to 32 bits
                    // and mask to be 32-byte-aligned.
addq %r15, %rax     // Add %r15, the sandbox base address.
jmp *%rax           // Indirect jump.
```

Exploiting the DRAM rowhammer bug to gain kernel privileges (Seaborn and Dullien)

Native Client (NaCl) Sandbox Escape

We can flip bits to allow for (unsafe) non 32-byte-aligned jumps!

Exploited "Safe" Code:

```
andl $~31, %eax ecx // Truncate address to 32 bits
                    // and mask to be 32-byte-aligned.
addq %r15, %rax     // Add %r15, the sandbox base address.
jmp *%rax           // Indirect jump.
```

Exploiting the DRAM rowhammer bug to gain kernel privileges (Seaborn and Dullien)

Kernel Privilege Escalation

- What could happen if a user could gain direct write access to a page table?

Each Page Table Entry (PTE) is 64 bits, containing:

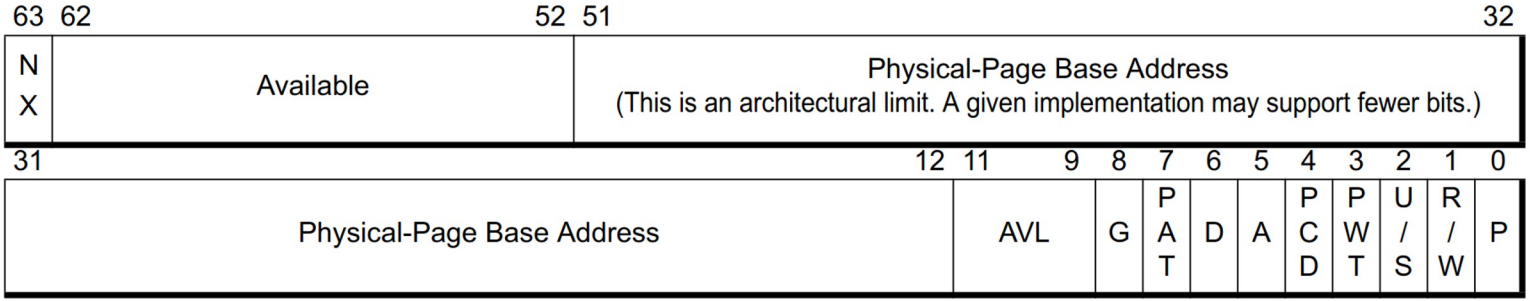
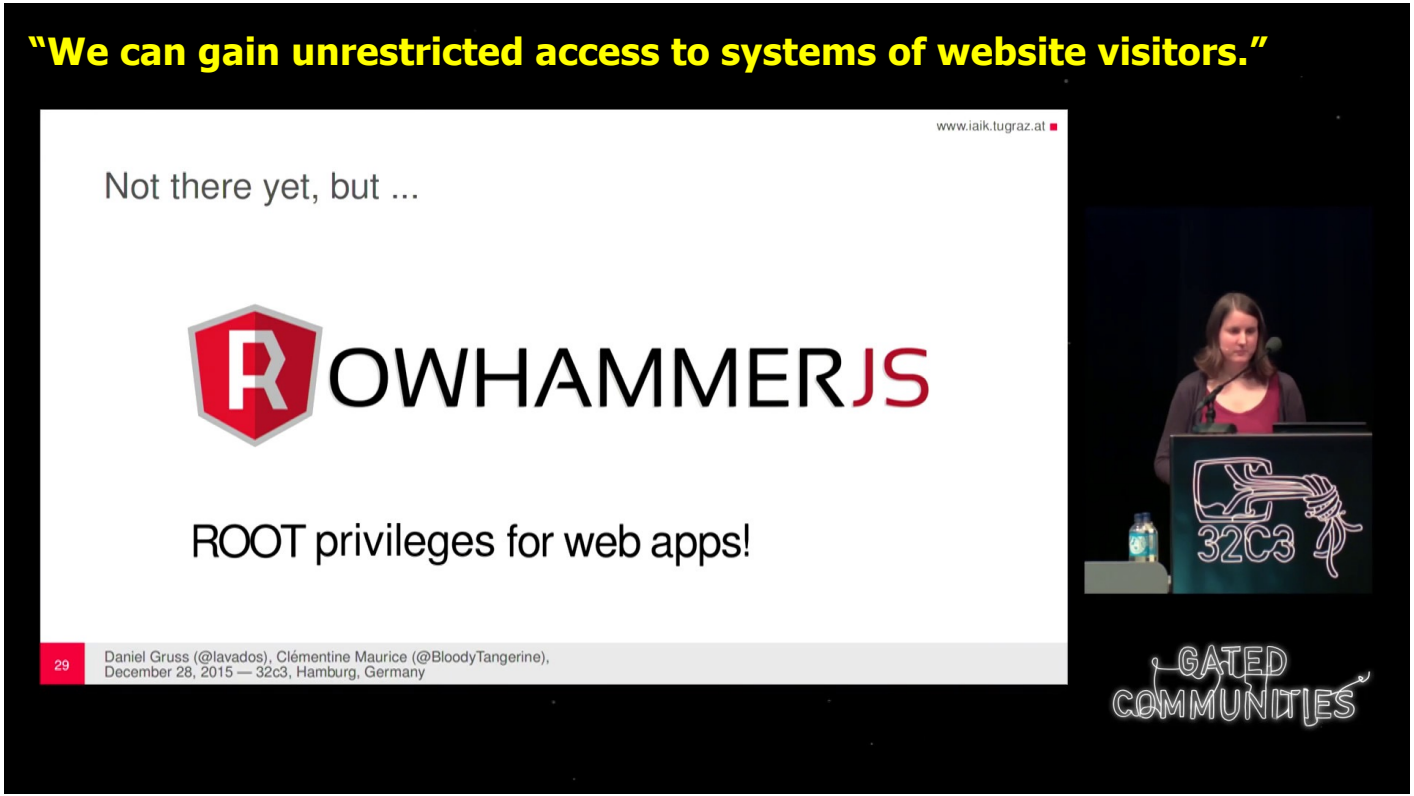


Figure 5-21. 4-Kbyte PTE—Long Mode

More Security Implications (I)



Rowhammer.js: A Remote Software-Induced Fault Attack in JavaScript (DIMVA'16)

Source: https://lab.dsst.io/32c3-slides/7197.html

More Security Implications (II)

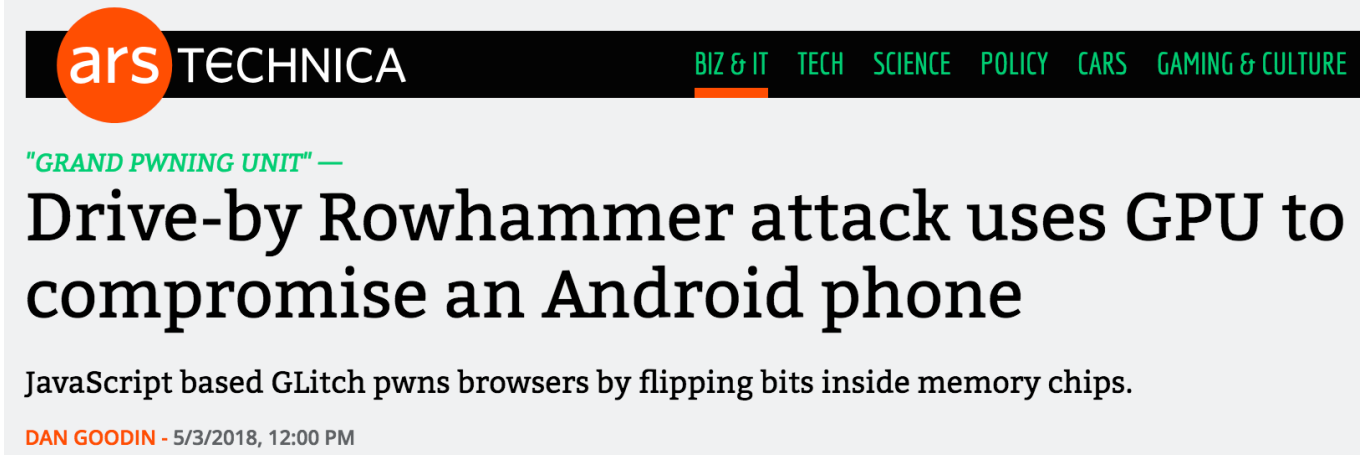
"Can gain control of a smart phone deterministically"

Drammer: Deterministic Rowhammer Attacks on Mobile Platforms, CCS'16

Source: https://fossbytes.com/drammer-rowhammer-attack-android-root-devices/

More Security Implications (III)

- Using an integrated GPU in a mobile system to remotely escalate privilege via the WebGL interface. IEEE S&P 2018



Grand Pwning Unit: Accelerating Microarchitectural Attacks with the GPU

Pietro Frigo
Vrije Universiteit
Amsterdam
p.frigo@vu.nl

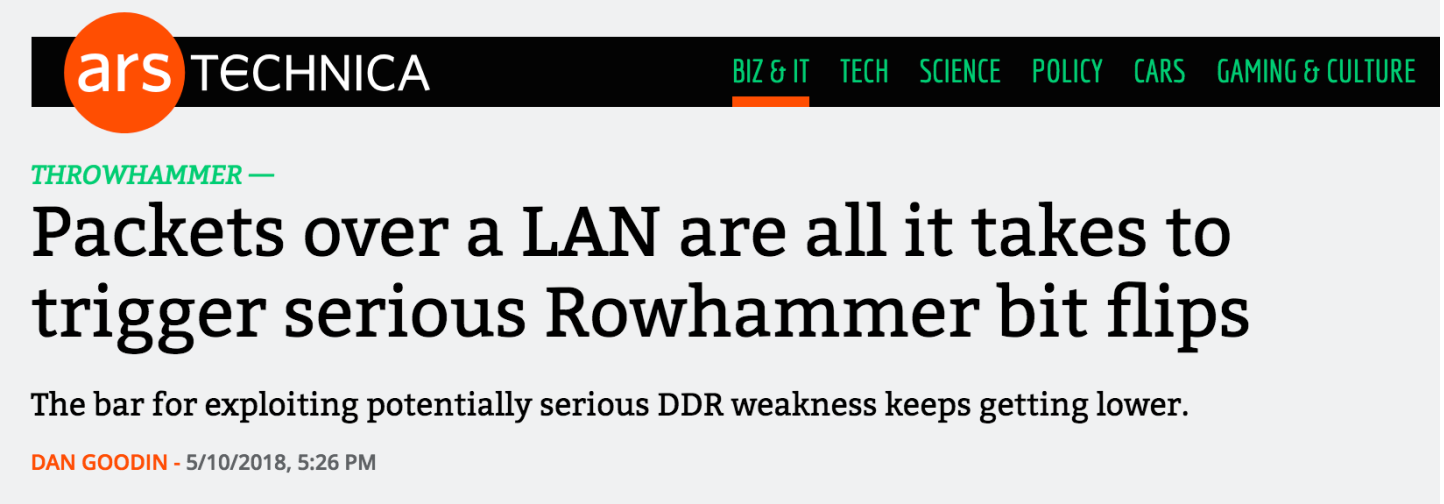
Cristiano Giuffrida
Vrije Universiteit
Amsterdam
giuffrida@cs.vu.nl

Herbert Bos
Vrije Universiteit
Amsterdam
herbertb@cs.vu.nl

Kaveh Razavi
Vrije Universiteit
Amsterdam
kaveh@cs.vu.nl

More Security Implications (IV)

- Rowhammer over RDMA (I) USENIX ATC 2018



ars TECHNICA BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULTURE

THROWHAMMER —

Packets over a LAN are all it takes to trigger serious Rowhammer bit flips

The bar for exploiting potentially serious DDR weakness keeps getting lower.

DAN GOODIN - 5/10/2018, 5:26 PM

Throwhammer: Rowhammer Attacks over the Network and Defenses

Andrei Tatar, *VU Amsterdam*
Radhesh Krishnan, *VU Amsterdam*
Elias Athanasopoulos, *University of Cyprus*
Cristiano Giuffrida, *VU Amsterdam*
Herbert Bos, *VU Amsterdam*
Kaveh Razavi, *VU Amsterdam*

More Security Implications (V)

- Rowhammer over RDMA (II)

The Hacker News™
Security in a serious way

Nethammer—Exploiting DRAM Rowhammer Bug Through Network Requests

Nethammer:
Inducing Rowhammer Faults through Network Requests

Moritz Lipp
Graz University of Technology

Misiker Tadesse Aga
University of Michigan

Michael Schwarz
Graz University of Technology

Daniel Gruss
Graz University of Technology

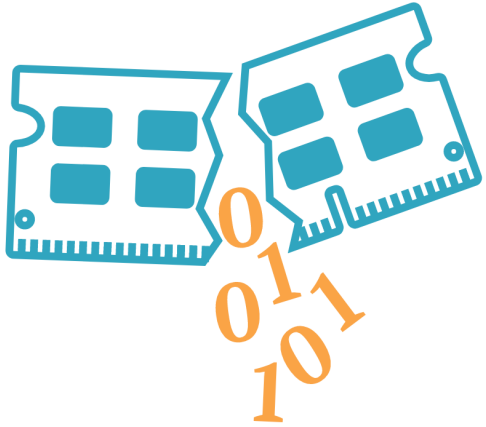
Clémentine Maurice
Univ Rennes, CNRS, IRISA

Lukas Raab
Graz University of Technology

Lukas Lamster
Graz University of Technology

More Security Implications (VI)

- IEEE S&P 2020



RAMBleed

RAMBleed: Reading Bits in Memory Without Accessing Them

Andrew Kwong
University of Michigan
ankwong@umich.edu

Daniel Genkin
University of Michigan
genkin@umich.edu

Daniel Gruss
Graz University of Technology
daniel.gruss@iaik.tugraz.at

Yuval Yarom
University of Adelaide and Data61
yval@cs.adelaide.edu.au

More Security Implications (VII)

- USENIX Security 2019

Terminal Brain Damage: Exposing the Graceless Degradation in Deep Neural Networks Under Hardware Fault Attacks

Sanghyun Hong, Pietro Frigo†, Yiğitcan Kaya, Cristiano Giuffrida†, Tudor Dumitraş

University of Maryland, College Park
†Vrije Universiteit Amsterdam

A Single Bit-flip Can Cause Terminal Brain Damage to DNNs

One specific bit-flip in a DNN's representation leads to accuracy drop over 90%

Our research found that a specific bit-flip in a DNN's bitwise representation can cause the accuracy loss up to 90%, and the DNN has 40-50% parameters, on average, that can lead to the accuracy drop over 10% when individually subjected to such single bitwise corruptions...

Read More

More Security Implications (VIII)

- USENIX Security 2020

DeepHammer: Depleting the Intelligence of Deep Neural Networks through Targeted Chain of Bit Flips

Fan Yao
University of Central Florida
fan.yao@ucf.edu

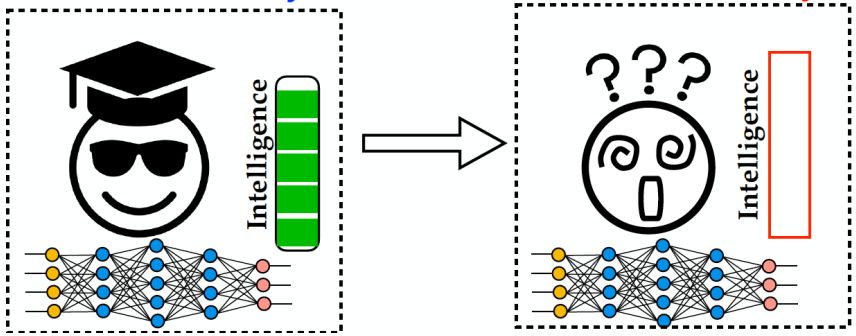
Adnan Siraj Rakin
Arizona State University
asrakin@asu.edu

Deliang Fan
dfan@asu.edu

Degrade the **inference accuracy** to the level of **Random Guess**

Example: ResNet-20 for CIFAR-10, **10** output classes

Before attack, **Accuracy: 90.2%** After attack, **Accuracy: ~10% (1/10)**



More Security Implications (IX)

- Rowhammer on MLC NAND Flash (based on [Cai+, HPCA 2017])

Security

Rowhammer RAM attack adapted to hit flash storage

Project Zero's two-year-old dog learns a new trick

By Richard Chirgwin 17 Aug 2017 at 04:27 17 SHARE ▼

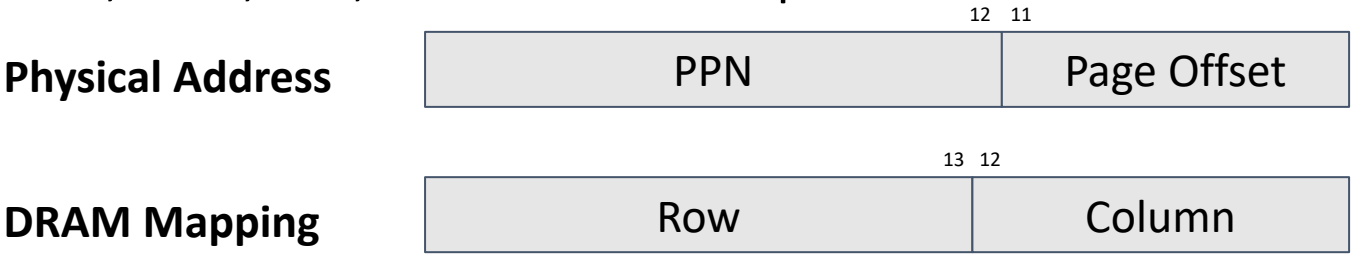
From random block corruption to privilege escalation: A filesystem attack vector for rowhammer-like attacks

Anil Kurmus Nikolas Ioannou Matthias Neugschwandtner Nikolaos Papandreou
Thomas Parnell
IBM Research – Zurich

More Security Implications?

Challenge: DRAM Addressing

- Different parts of a physical address are used to determine which rank, bank, row, and column it maps to



- Accesses to different banks allow for parallelism!

Bank ID = $[(13 \oplus 15), (14 \oplus 16), (17 \oplus 19)]$

Rowhammer Mitigations?

- Manufacturing “better” chips
- Increasing refresh rate
- Error Correcting Codes
- Targeted row refresh (TRR) - Used in DDR4!
- Retiring vulnerable cells
- Static binary analysis
- User/kernel space isolation in physical memory

Rowhammer Solutions?

- Manufacturing “better” chips — cost
- Increasing refresh rate — Performance, power
- Error Correcting Codes — cost, power
- Targeted row refresh (TRR) - Used in DDR4! — cost, power, complexity
- Retiring vulnerable cells — cost, power, complexity
- Static binary analysis — security
- User/kernel space isolation in physical memory

Takeaways

- DRAM basic structures

- What is Rowhammer attacks?
  - What is the trend forward?
  - What attacks can you build with Rowhammer as a primitive?
  - How to mitigate Rowhammer attacks?