Memory Systems
and Memory-Centric Computing Systems
Lecture 1b: Course Logistics

Prof. Onur Mutlu
omutlu@gmail.com
https://people.inf.ethz.ch/omutlu

12 June 2019

TU Wien Fast Course 2019

SAFARI ETH zürich Carnegie Mellon

Course Logistics

What Will You Learn in This Course?

- Memory Systems and Memory-Centric Computing Systems
  - June 12-19, 2019

- Topic 1: Main Memory Trends and Fundamentals
- Topic 2: Memory Reliability & Security: RowHammer and Beyond
- Topic 3: In-memory Computation
- Topic 4: Low-Latency and Low-Energy Memory
- Topic 5: Enabling and Exploiting Emerging Memory Technologies
- Topic 6: Memory Interference and QoS
- Topic 7: Energy-efficient Interconnects
- Topic 8: Flash Memory and SSD Scaling

This Course

- Will cover many problems and potential solutions related to the design of memory systems in the many core era

- The design of the memory system poses many
  - Difficult research and engineering problems
  - Important fundamental problems
  - Industry-relevant problems
  - **Problems whose solutions can revolutionize the world**

- Many creative and insightful solutions are needed to solve these problems

- Goal: Acquire the basics to develop such solutions (by covering fundamentals and cutting edge research)

Course Information

- My Contact Information
  - Onur Mutlu
  - omutlu@gmail.com
  - https://people.inf.ethz.ch/omutlu
  - +41-79-572-1444 (my cell phone, Whatsapp)
  - Find me during breaks and/or email any time.

- Website for Course Slides, Papers, Updates
  - https://safari.ethz.ch/memory_systems/TUWien2019/

- For the curious:
  - ACACES 2018: Memory Systems and Memory-Centric Computing Systems
  - https://safari.ethz.ch/memory_systems/ACACES2018/doku.php

Course Staff

- Onur Mutlu
  - Instructor
  - omutlu@gmail.com

- Abdullah Hanif
  - TA

How To Make the Best Out of This Course

- Be alert during lectures – they will be fast paced
- Do the readings (and explore even more)
  - I will provide many references + review assignments
- Go back and reinforce fundamentals (as needed)
  - I will provide pointers to basic computer architecture materials (lecture videos, slides, readings, exams, …)

- Remember "Chance favors the prepared mind." (Pasteur)

How to Do a Good Paper Review

Anatomy of a Good Paper Review (Talk)

- 1: Summary
  - What is the problem the paper is trying to solve?
  - What are the key ideas of the paper? Key insights?
  - What are the key mechanisms? What is the implementation?
  - What are the key results? Key conclusions?
- 2: Strengths (most important ones)
  - Does the paper solve the problem well? Is it well written? …
- 3: Weaknesses (most important ones)
  - This is where you should think critically. Every paper/idea has a weakness. This does not mean the paper is necessarily bad. It means there is room for improvement and future research can accomplish this.
- 4: Thoughts/Ideas: Can you do better? Present your ideas.
- 5: Takeaways: What you learned/enjoyed/disliked? Why?
- 6: Discussion starters and questions.

- Review should be short and concise (20 minutes or < one page)

Suggested Paper Review Format

- Problem & Goal
- Key Ideas/solution
- Novelty
- Mechanisms & Implementation
- Major Results
- Takeaways/Conclusions

Summary

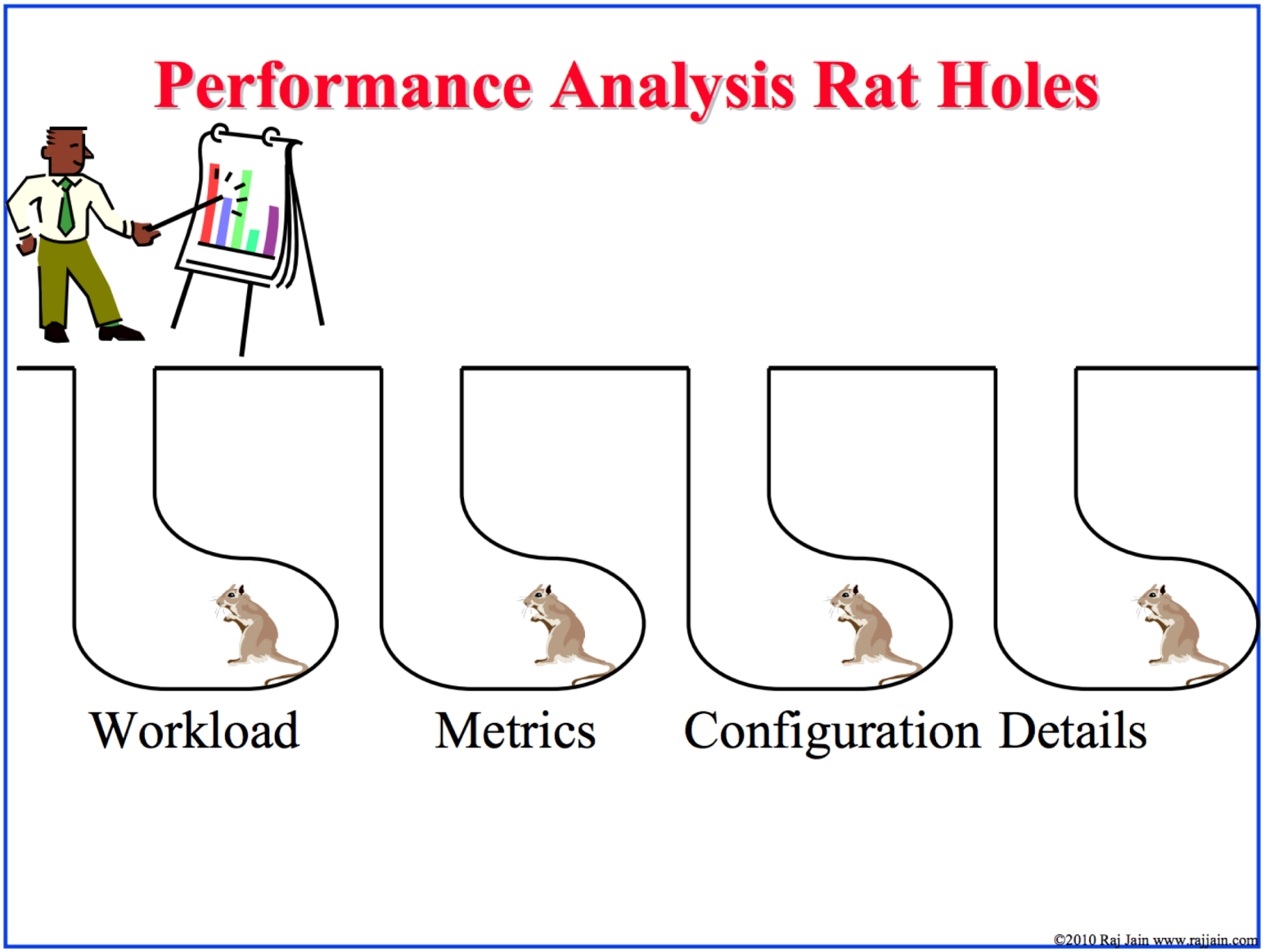
- Strengths
- Weaknesses
- Alternatives
- New ideas/problems
- Brainstorming and Discussion

Critique & New Ideas

More Advice on Paper Review Talk

- When doing the paper reviews and analyses, be very critical
- Always think about better ways of solving the problem or related problems
  - Question the problem as well
  - Read background papers (both past and future)
- This is how things progress in science and engineering (or anywhere), and how you can make big leaps
  - By critical analysis
- A few sample text reviews will be provided online

Try to Avoid Rat Hole Discussions



Source: https://www.cse.wustl.edu/~jain/iucee/ftp/k_10adp.pdf

10.8 DECISION MAKER'S GAMES

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

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

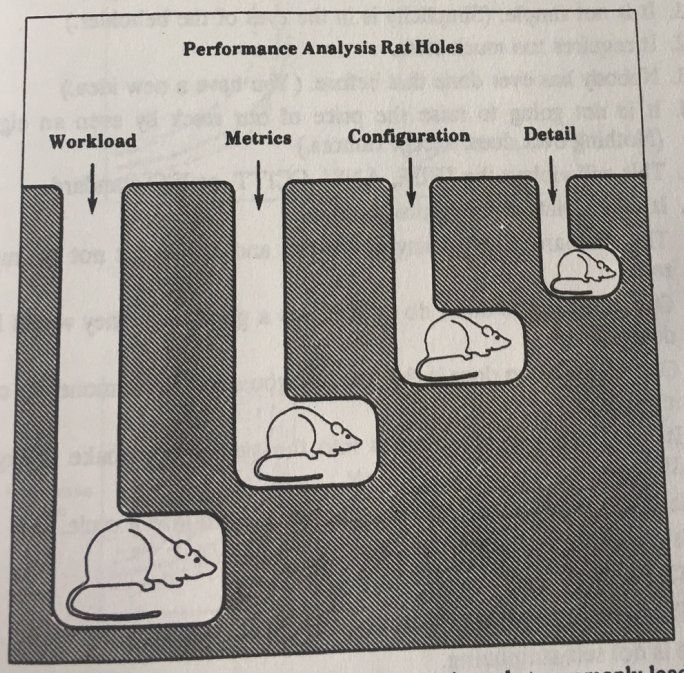


FIGURE 10.26 Four issues in performance presentations that commonly lead to endless discussion.

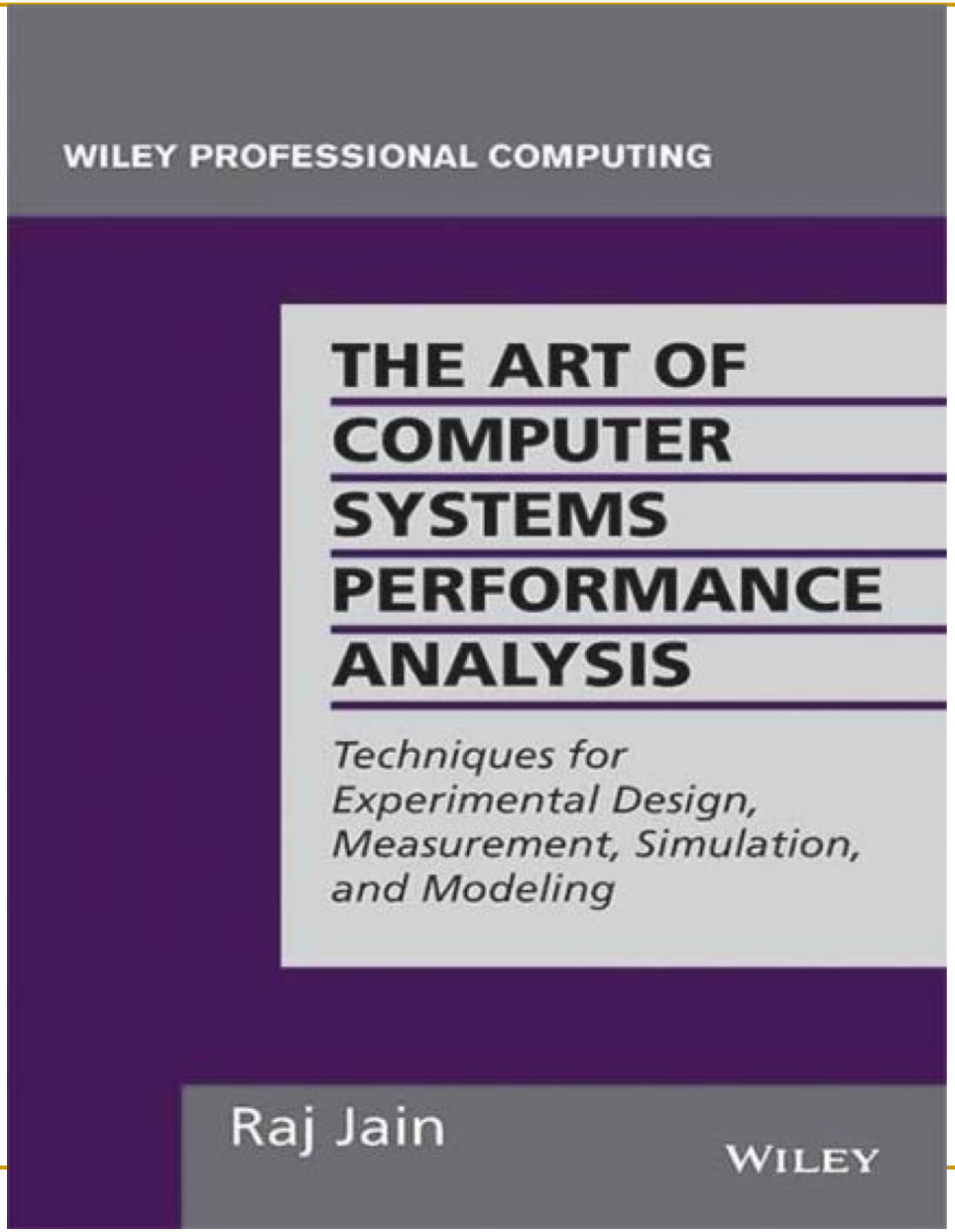
Raj Jain, "The Art of Computer Systems Performance Analysis," Wiley, 1991.

Box 10.2 Reasons for Not Accepting the Results of an Analysis

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

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

Aside: A Recommended Book



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

More Advice on Talks

- Kayvon Fatahalian, "Tips for Giving Clear Talks"
  - http://graphics.stanford.edu/~kayvonf/misc/cleartalktips.pdf
  - Many useful and simple principles here

"Every sentence matters"

"The audience prefers not to think" (about things you can just tell them)

"Surprises are bad": say why before what
(indicate why you are saying something before you say it)

Explain every figure, graph, or equation

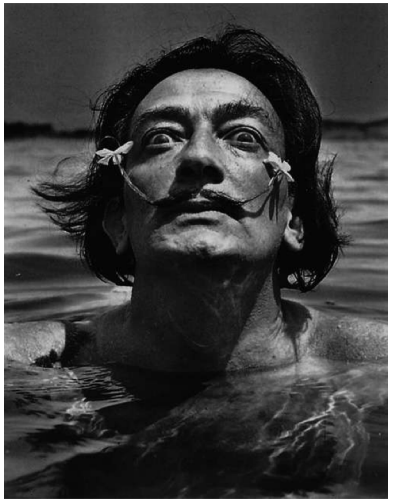
When improving the talk, the audience is always right

Who Painted This Painting?

Salvador Dali @ 1924

Source: http://www.museoreinasofia.es/en/collection/artwork/retrato-luis-bunuel-portrait-luis-bunuel http://salvadordaliprints.org/biography/

What About This?



Salvador Dali @ 1937

Source: http://salvadordaliprints.org/the-burning-giraffe/ https://www.dalipaintings.com/

Takeaway

Learn the basic principles

before you

consciously choose to break them

Readings, Videos, Reference Materials

Accelerated Memory Course (~6.5 hours)

- ACACES 2018
  - Memory Systems and Memory-Centric Computing Systems
  - Taught by Onur Mutlu July 9-13, 2018
  - ~6.5 hours of lectures

- Website for the Course including Videos, Slides, Papers
  - https://safari.ethz.ch/memory_systems/ACACES2018/
  - https://www.youtube.com/playlist?list=PL5Q2soXY2Zi-HXxomthrpDpMJm05P6J9x

- All Papers are at:
  - https://people.inf.ethz.ch/omutlu/projects.htm
  - Final lecture notes and readings (for all topics)

Required: Reference Overview Paper I

Processing Data Where It Makes Sense: Enabling In-Memory Computation

Onur Mutlu[a,b], Saugata Ghose[b], Juan Gómez-Luna[a], Rachata Ausavarungnirun[b,c]

[a] *ETH Zürich*
[b] *Carnegie Mellon University*
[c] *King Mongkut's University of Technology North Bangkok*

Onur Mutlu, Saugata Ghose, Juan Gomez-Luna, and Rachata Ausavarungnirun,
"Processing Data Where It Makes Sense: Enabling In-Memory Computation"
*Invited paper in Microprocessors and Microsystems (**MICPRO**)*, June 2019.
[arXiv version]

Required: Reference Overview Paper II

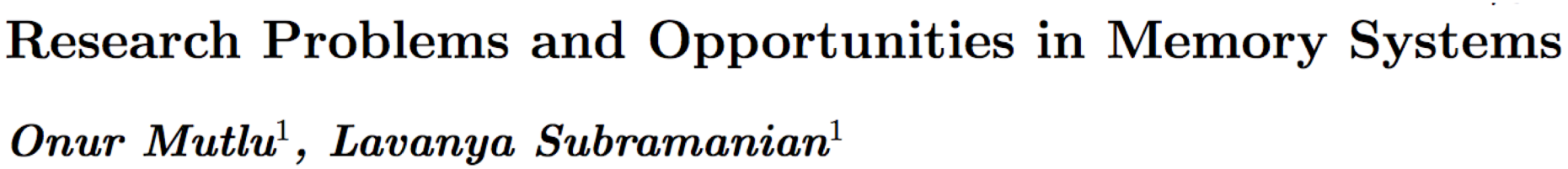
- Onur Mutlu and Jeremie Kim,
  "RowHammer: A Retrospective"
  IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (***TCAD***) *Special Issue on Top Picks in Hardware and Embedded Security*, 2019.
  [Preliminary arXiv version]

RowHammer: A Retrospective

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

Required: Reference Overview Paper III

- Onur Mutlu and Lavanya Subramanian,
  "Research Problems and Opportunities in Memory Systems"
  *Invited Article in Supercomputing Frontiers and Innovations (**SUPERFRI**)*, 2014/2015.


Research Problems and Opportunities in Memory Systems

Onur Mutlu[1], *Lavanya Subramanian*[1]

https://people.inf.ethz.ch/omutlu/pub/memory-systems-research_superfri14.pdf

Reference Overview Paper IV

Enabling the Adoption of Processing-in-Memory: Challenges, Mechanisms, Future Research Directions

SAUGATA GHOSE, KEVIN HSIEH, AMIRALI BOROUMAND,
RACHATA AUSAVARUNGNIRUN
Carnegie Mellon University

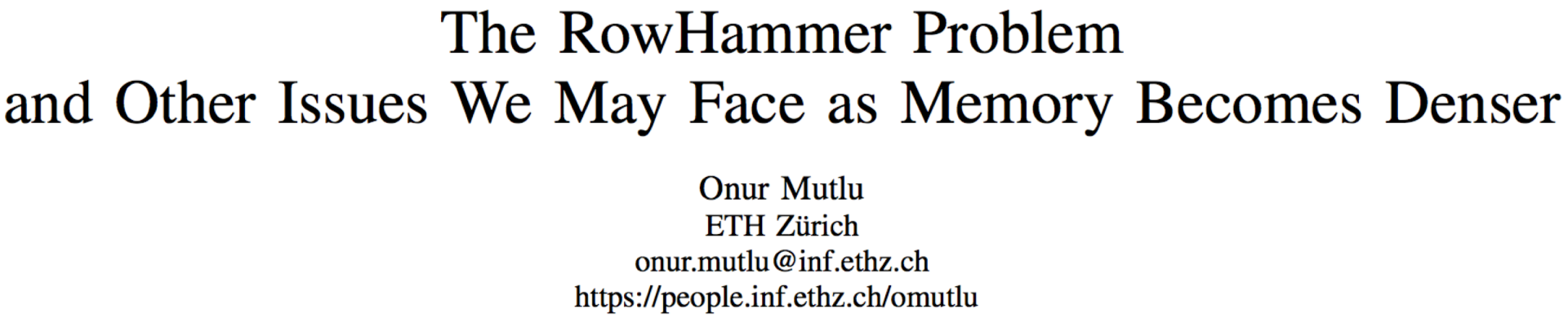
ONUR MUTLU
ETH Zürich and Carnegie Mellon University

Saugata Ghose, Kevin Hsieh, Amirali Boroumand, Rachata Ausavarungnirun, Onur Mutlu, **"Enabling the Adoption of Processing-in-Memory: Challenges, Mechanisms, Future Research Directions"**
Invited Book Chapter, to appear in 2018.
[Preliminary arxiv.org version]

https://arxiv.org/pdf/1802.00320.pdf

Reference Overview Paper V

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



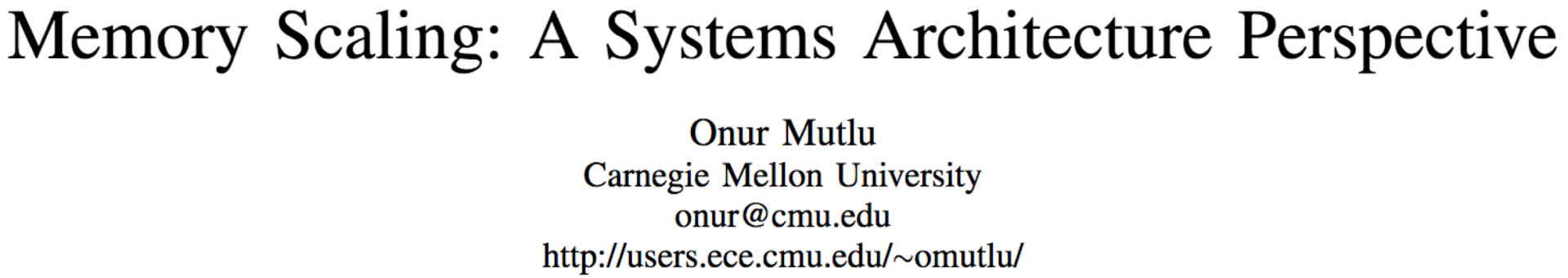
The RowHammer Problem
and Other Issues We May Face as Memory Becomes Denser

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

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

Reference Overview Paper VI

- Onur Mutlu,
  "Memory Scaling: A Systems Architecture Perspective"
  *Technical talk at MemCon 2013 (**MEMCON**)*, Santa Clara, CA, August 2013. [Slides (pptx) (pdf)]
  [Video] [Coverage on StorageSearch]



Memory Scaling: A Systems Architecture Perspective

Onur Mutlu
Carnegie Mellon University
onur@cmu.edu
http://users.ece.cmu.edu/~omutlu/

https://people.inf.ethz.ch/omutlu/pub/memory-scaling_memcon13.pdf

Reference Overview Paper VII

Proceedings of the IEEE, Sept. 2017

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

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

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

https://arxiv.org/pdf/1706.08642

Related Videos and Course Materials (I)

- Undergraduate Computer Architecture Course Lecture Videos (2015, 2014, 2013)
- Undergraduate Computer Architecture Course Materials (2015, 2014, 2013)

- Graduate Computer Architecture Course Lecture Videos (2017, 2015, 2013)
- Graduate Computer Architecture Course Materials (2017, 2015, 2013)

- Parallel Computer Architecture Course Materials (Lecture Videos)

Related Videos and Course Materials (II)

- **Freshman Digital Circuits and Computer Architecture Course Lecture Videos (2018, 2017)**
- **Freshman Digital Circuits and Computer Architecture Course Materials (2018)**

- **Memory Systems Short Course Materials**
  (Lecture Video on Main Memory and DRAM Basics)

Some Open Source Tools (I)

- Rowhammer – Program to Induce RowHammer Errors
  - https://github.com/CMU-SAFARI/rowhammer
- Ramulator – Fast and Extensible DRAM Simulator
  - https://github.com/CMU-SAFARI/ramulator
- MemSim – Simple Memory Simulator
  - https://github.com/CMU-SAFARI/memsim
- NOCulator – Flexible Network-on-Chip Simulator
  - https://github.com/CMU-SAFARI/NOCulator
- SoftMC – FPGA-Based DRAM Testing Infrastructure
  - https://github.com/CMU-SAFARI/SoftMC
- Other open-source software from my group
  - **https://github.com/CMU-SAFARI/**
  - **http://www.ece.cmu.edu/~safari/tools.html**

Some Open Source Tools (II)

- MQSim – A Fast Modern SSD Simulator
  - https://github.com/CMU-SAFARI/MQSim
- Mosaic – GPU Simulator Supporting Concurrent Applications
  - https://github.com/CMU-SAFARI/Mosaic
- IMPICA – Processing in 3D-Stacked Memory Simulator
  - https://github.com/CMU-SAFARI/IMPICA
- SMLA – Detailed 3D-Stacked Memory Simulator
  - https://github.com/CMU-SAFARI/SMLA
- HWASim – Simulator for Heterogeneous CPU-HWA Systems
  - https://github.com/CMU-SAFARI/HWASim
- Other open-source software from my group
  - **https://github.com/CMU-SAFARI/**
  - **http://www.ece.cmu.edu/~safari/tools.html**

More Open Source Tools (III)

- A lot more open-source software from my group
  - **https://github.com/CMU-SAFARI/**
  - **http://www.ece.cmu.edu/~safari/tools.html**

Referenced Papers

- All are available at

  https://people.inf.ethz.ch/omutlu/projects.htm

  http://scholar.google.com/citations?user=7XyGUGkAAAAJ&hl=en

  https://people.inf.ethz.ch/omutlu/acaces2018.html

Grading

Grading

- At least 8 paper reviews
  - 12 recommended
  - We will announce how and when to submit the reviews
- Various assignments given during lecture
- Participation important

Your First Assignment

- Due tonight, June 12, 23:59

- Please send me an email at omutlu@gmail.com
  - Subject: TU Wien Memory Systems June 2019 <your name>
  - Introduce yourself and answer several questions
  - What is your educational background?
  - What degree you are studying for?
  - Why are you taking this course?
  - What is your research topic?
  - What are your career goals?
  - What is your passion?
  - Any other thing you would like to tell me…

Memory Systems
and Memory-Centric Computing Systems
Lecture 1b: Course Logistics

Prof. Onur Mutlu
omutlu@gmail.com
https://people.inf.ethz.ch/omutlu

12 June 2019

TU Wien Fast Course 2019

SAFARI ETH zürich Carnegie Mellon