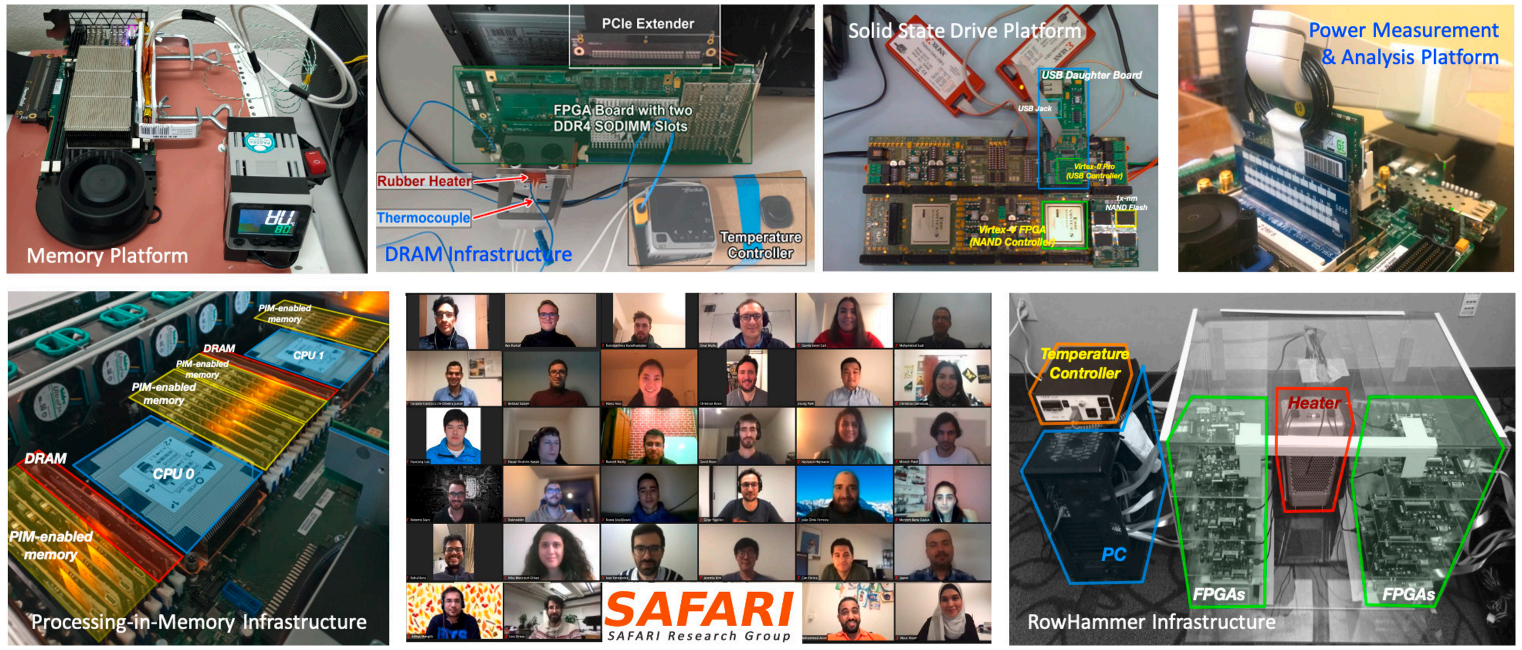

SAFARI Research Group
Professor Onur Mutlu

The SAFARI Research Group conducts cutting-edge research and education in computer architecture, computing systems, hardware security, bioinformatics, and software/hardware co-design. The major goal is to design fundamentally better computing architectures. Our work spans the boundaries between applications, systems, languages, system software, compilers, and hardware, with architecture at the core. We tackle many fundamental issues in high performance, energy efficiency, hardware security, fault tolerance, predictable systems, dependable systems, hardware/software cooperation, and genome analysis. We are especially excited about novel, fundamentally secure and fundamentally efficient computation, communication and memory/storage paradigms. Recent impactful examples of the group's work include leading research into memory/storage systems and new computing paradigms, e.g. the discovery of the DRAM RowHammer vulnerability and the development of a comprehensive Processing-in-Memory paradigm.


Prof. Onur Mutlu

Curriculum vitae

Degrees / higher education

2006 PhD in Computer Engineering, University of Texas at Austin, USA
2002 MSE in Computer Engineering, University of Texas at Austin, USA
2000 BSE in Computer Engineering, BS in Psychology, with Highest Distinction, University of Michigan, Ann Arbor, USA

Professional experience

since 2015 Full Professor, ETH Zurich
since 2016 Carnegie Mellon University, Adjunct Professor
2009–2016 Carnegie Mellon University, Strecker Early Career Professor
2016 Google & VMware, Visiting Research Scientist
since 2015 Bilkent University, Adjunct Professor
2012 Intel, Visiting Researcher
2006–2009 Microsoft Research, Researcher
2001–2005 Intel & AMD, Coop Engineer

Top picture: Hardware infrastructures developed and open-sourced by the group: (top) Memory, DRAM, SSD, power infrastructures; (bottom) Processing-In-Memory infrastructure; group picture; RowHammer infrastructure; more at https://safari.ethz.ch/

Group structure and resources

Staff	FTE*
Professor(s)	1
Senior scientists	–
Oberassistent	1
Postdoc	6
Doctoral students	11
Semester/Master's students	4
Technical staff	–
Administrative staff	1
Apprentices	–

(*as per 31 Dec. 2020)

Space	
Labs	x m²
Office	x m²
Other	x m²

Teaching hours 2020		
Semester	spring	fall
No. of courses	2	6
Hours/week	8	21

Funding	
ETH Zurich core funding	x
ETH Zurich competitive funding	x
Funding national	x
Funding international	x
Industry and foundations	x

Annually, average 2016 – 2021

Doctoral graduations					
Year	2020	2019	2018	2017	2016
Graduations	2	3	2	2	3

Active collaborations				
Institution	D-ITET	ETH	National	International
No. of groups	2	2	2	50+

Goals and priorities

The overarching goal of the SAFARI Research Group is to design fundamentally better computer architectures that will have long-lasting positive impact on human lives. We aim to invent computer architecture paradigms that get widely used for decades to come. The group develops both theoretical and practical ideas, methods, and infrastructures to understand, model and enhance performance, energy efficiency, reliability, security, safety, predictability, dependability, usability, and scalability of computer architectures. We have an extensive research and education portfolio centred around computer architecture and systems, with significant concerted efforts in all aspects of memory and storage, hardware security/safety/reliability, and system design for bioinformatics, medicine, and machine learning (with special focus on genome analysis).

Memory (and data movement from memory) is a key bottleneck in all computing systems. We tackle key problems in computer memory systems, spanning circuits to algorithms/applications. We rigorously study how to make memory systems fundamentally secure/safe/reliable, low-latency, predictable, and energy-efficient and how to customize them for emerging applications like genomics, artificial intelligence, medical/health systems. Solutions to memory technology scaling problems, data-centric computing paradigms (Processing-in-Memory), and algorithm-architecture co-design for fast and efficient genome analysis are three major directions.
The group has strong collaborations with industry and an extensive open-source teaching portfolio at BS/MS/PhD levels.

Research impact

Data storage and movement is a fundamental bottleneck in modern computing systems and applications, for performance, energy efficiency, predictability, reliability, security, and safety. Our research has had significant impact in scientifically understanding and mitigating this bottleneck, with many ideas impacting industrial products and igniting new research areas. We expect and hope this impact to continue and become larger since the memory bottleneck and data are only becoming larger.

One key example is our discovery and analyses of the RowHammer problem, the first hardware failure mechanism that causes a practical and widespread system security vulnerability. RowHammer heavily influenced industrial products (of memory manufacturers, processor manufacturers, and system vendors) and academic research in security and hardware.

Teaching activities

Our teaching portfolio includes Bachelor's-level (500+-student) "Digital Design & Computer Architecture", Master's-level (30–50-student) "Computer Architecture", as well as seminar (20+-student) and hands-on project (3–10-student) courses in computer architecture, genome analysis, processing in memory, heterogeneous systems, DRAM, and solid-state drives, some of which are co-led by our senior scientists. We are a strong advocate for free and open sharing of teaching and research artifacts to democratize education worldwide. Our course videos/materials are freely available online, and used by many educators, researchers and practitioners, including leading companies. Many students and universities without access to state-of-the-art computer architecture classes benefit from our online classes (see http://youtube.com/onurmutlulectures).

Promotion of young academics

Since 2015, we graduated 14 PhD students and five postdocs. Two PhDs received prestigious dissertation awards (EDAA Outstanding Dissertation; IEEE Turkey Best PhD Thesis). Three PhDs and three postdocs are faculty members at prestigious universities (e.g. Univ. of Toronto, Illinois, Virginia). Eight PhDs are research scientists at prestigious industrial labs (e.g. Microsoft Research, NVIDIA Research, Google, Facebook, Intel). Several PhDs and postdocs won prestigious fellowships (e.g. Microsoft, Intel, Google, Facebook, NVIDIA, SRC, Qualcomm PhD fellowships and Korean Science Foundation Postdoc/PhD fellowships). We promote an open, collaborative, communicative scientific culture that aims to enable every group member to exceed their perceived potential through enablement, empowerment, and collaboration.

Research activities and achievements

A New memory and storage architectures

DRAM (Dynamic Random Access Memory) is the predominant technology used for computer memory. It is facing significant challenges in technology scaling, reliability, data retention, latency, bandwidth, and power consumption. These greatly affect performance, energy, security/safety/reliability and scalability of computing platforms and applications. We rigorously understand and solve these challenges via novel techniques across the computing stack. To this end, we build hardware infrastructures (see top picture) and follow two key directions.

A.1 Fundamentally better DRAM architectures

We experimentally demonstrated, analyzed and proposed solutions for the RowHammer problem that affects most modern DRAM chips. We were the first to show that, by repeatedly accessing a DRAM row, one can induce errors in adjacent rows. A malicious attacker can use this to circumvent memory protection and gain complete control over an otherwise secure system. RowHammer is the first example of a device failure mechanism that causes a practical, widespread system security vulnerability. Our RowHammer work (ISCA'14'20, S&P'20, HPCA'21) continues to have widespread impact on security and hardware communities. For example, our work led to inclusion of new tests in widely used memtest programmes; Apple cited our work in security releases; Intel and other vendors implemented our major solutions; our 2020 works led to industry-wide task groups to solve RowHammer and Best Paper Awards.

We pioneer research on solving critical DRAM scaling problems (refresh, latency, variability, power, energy, reliability) by analyzing real chips, improving DRAM in all directions with large impact. For example, Intel & Samsung advocated several of our ideas for future DRAM standards. Our work to eliminate memory refresh and data retention failures influenced academic and industrial directions (e.g. works on DRAM Error Correcting Codes in DSN'19 and MICRO'20 won Best Paper Awards).

A.2 Enabling emerging memory technologies

Emerging technologies, e.g. Phase Change Memory, magnetic memory and memristors, have promising properties as memory / storage devices but also large downsides. We do research to enable and exploit such technologies by designing intelligent architectures. Our work served as a precursor of Intel's Optane Memory and other technologies being designed for hybrid memory.

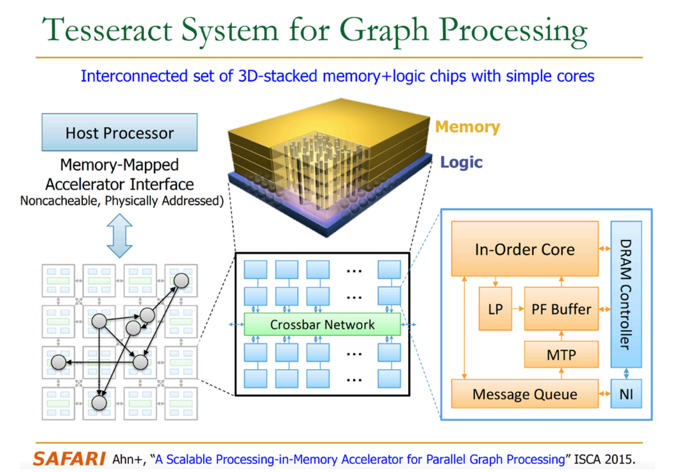

Figure 1: Our Tesseract Processing-in-Memory system for graph processing (ISCA'15) provides more than 13 × performance improvement and 8 × energy reduction over state-of-the-art systems. Many works have been built on Tesseract, which provides a blueprint for future PIM systems.

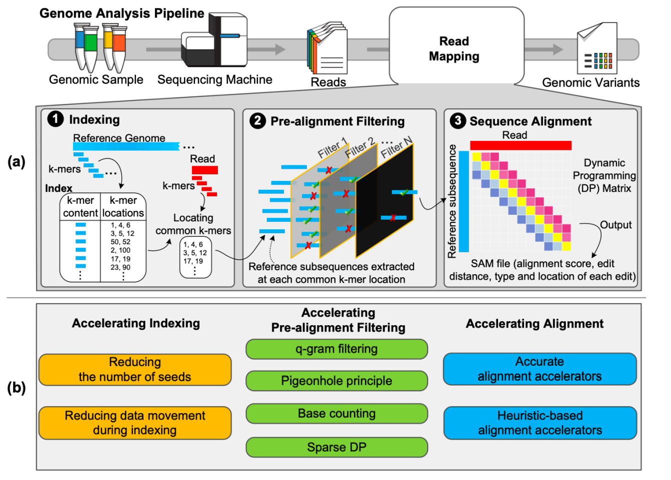

Figure 2: Our bioinformatics work covers the entire genome analysis pipeline. Collectively, our algorithm-architecture co-design techniques provide >100 × performance improvement and energy reduction over state-of-the-art systems. Figure replicated from our IEEE Micro 2020 invited paper "Accelerating Genome Analysis: A Primer on an Ongoing Journey".

B Data-centric paradigms: Processing-in-Memory

Modern computing systems are processor-centric, i.e. overwhelmingly designed to move data to computation. This greatly exacerbates performance, energy and scalability problems because data movement is orders of magnitude more costly than computation (in latency and energy).

We do research to fundamentally change the design paradigm of computers: to enable computation near data, i.e. Processing-in-Memory (PIM). PIM places computation in/near where data is stored (inside memory chips, in logic layers of 3D-stacked memory, in memory controllers), so that data movement is reduced. We develop at least two new approaches to PIM: 1) processing using memory by exploiting analogue operational properties of memory to perform massively-parallel processing, 2) processing near memory by exploiting memory-logic integration technologies to provide near-memory logic.

We pioneer modern PIM research, tackling all aspects of how to enable and design PIM systems: cross-layer research, design, and adoption challenges in devices, architecture, systems, applications and programming models. Our work has influenced academia and industry (see figure 1). We work closely with industry (e.g. UPMEM, Google, Microsoft, Facebook, Intel, ASML, SRC) to enable adoption of our PIM paradigms.

C Fast and efficient genome analysis, medicine, ML / AI

Genome analysis is the foundation of many scientific and medical discoveries, and a key enabler of personalized medicine. Current systems are too slow and too energy-inefficient. Our goal is to design fundamentally better genome analysis systems, enabling decisions within seconds/minutes (vs. days/weeks), using minimal energy. Such systems can revolutionize medicine, public health and scientific discovery. To this end, we develop novel algorithms and architectures, e.g. for DNA readmapping [NatureGenetics'09, Bioinformatics'15'17'19'20] and approximate string matching [MICRO'20]; (see figure 2). We also examine other dimensions, e.g. privacy, security and mobile/embedded genomics.

Collaborations

We believe strong collaboration with academic and industrial groups is critical for larger ideas and higher impact. We collaborate with various groups at ETH and many groups in the world (e.g. MIT, CMU, Illinois, Michigan, VU Amsterdam, TU Eindhoven, TU Wien, Seoul Nat'l Univ., NTU Athens, Bilkent). We have strong connections with industrial partners who provide donations for our research directions (e.g. Google, Huawei, Intel, Microsoft, VMware) and work closely with Semiconductor Research Corporation.

Participation in / contributions to technology platforms

We design and maintain unique cutting-edge hardware infrastructures that enable testing of memory chips (DRAM and flash memory) for reliability, security and performance issues, under various operating conditions (voltage, temperature, workloads) - see top picture. We also design and maintain architecture-level simulation tools (e.g. Ramulator), benchmark suites, and bioinformatics software (e.g. GateKeeper, SneakySnake). We open-source these artifacts (https://github.com/CMU-SAFARI), which are used by both academia and industry and led to many independent validations and new ideas published in literature.

Strengths, weaknesses and outlook

Strengths

SAFARI has a collaborative, constructive idea- and creativity-driven culture. We have developed impactful leading ideas in memory systems, computer architecture and genome analysis. We have strong ties with leading industrial and academic partners. Our open online courses have a strong following, amplifying our scientific activities and attracting strong excited students. We aim to continue to build on these strengths to take our research and educational activities to the next levels to maximize our impact.

Weaknesses

Building up our collaborations in D-ITET is a priority; we see large collaboration potential with many groups working on, e.g. devices, circuits, biomedical and vision systems, machine learning. We actively work on efficiently scaling the group size to better tackle large scientific problems. A thriving local computer hardware industry in Switzerland/Europe would enhance our local impact.

Outlook

We will continue our leading research in all aforementioned areas and also leap into new areas, e.g. fast outbreak prediction via mobile genomics and novel biologically inspired computing paradigms.

Selected highlights and awards

2021	IEEE High Performance Computer Architecture Test of Time Award
2020	IEEE Computer Society Edward J. McCluskey Technical Achievement Award
2020	Best Paper Awards @ MICRO and IEEE S&P + Pwnie Most Innovative Research Award
2020	EDAA Outstanding Dissertation Award
2019	ACM SIGARCH Maurice Wilkes Award
2019	DSN Best Paper + IEEE TCAD Top Pick Paper Award in Hardware Security
2019	Facebook and Google Faculty Research Awards
2018	IEEE Fellow and Academia Europaea Election
2018	IEEE Turkey Best PhD Dissertation Award
2017	ACM Fellow
2017	DFRWS-EU Best Paper Award
2014	RTAS Best Paper Award
2014	Google and Microsoft Faculty Research Awards
2013	Strecker Early Career Chair Professorship
2013	IBM Faculty Award
2012	Intel Early Career Faculty Award
2012	CMU Ladd Research Award
2012	ICCD Best Paper Award
2012	HP Labs and IBM Faculty Awards
2011	IEEE Computer Society Young Computer Architect Award
2011	11 IEEE Micro Top Pick Paper Awards since 2003
2010	Best Paper Awards @ ASPLOS and VTS
2010	NSF CAREER Award

10 Key publications

1 M. Patel et al., "Bit-Exact ECC Recovery (BEER): Determining DRAM On-Die ECC Functions by Exploiting DRAM Data Retention Characteristics", MICRO 2020. Best Paper Award.

2 D. Senol Cali et al., "GenASM: A High-Performance, Low-Power Approximate String Matching Acceleration Framework for Genome Sequence Analysis", MICRO 2020.

3 N. Hajinazar et al., "The Virtual Block Interface: A Flexible Alternative to the Conventional Virtual Memory Framework", ISCA 2020.

4 P. Frigo et al., "TRRespass: Exploiting the Many Sides of Target Row Refresh", S&P 2020. Best Paper Award; Pwnie Most Innovative Research Award

5 J.S. Kim et al., "D-RaNGe: Using Commodity DRAM Devices to Generate True Random Numbers with Low Latency and High Throughput", HPCA 2019. IEEE Micro Top Picks Honorable Mention.

6 A. Boroumand et al., "Google Workloads for Consumer Devices: Mitigating Data Movement Bottlenecks", ASPLOS 2018.

7 V. Seshadri et al., "Ambit: In-Memory Accelerator for Bulk Bitwise Operations Using Commodity DRAM Technology", MICRO 2017.

8 Y. Cai et al., "Error Characterization, Mitigation, and Recovery in Flash Memory Based Solid State Drives", Proceedings of the IEEE 2017.

9 J. Ahn et al., "A Scalable Processing-in-Memory Accelerator for Parallel Graph Processing", ISCA 2015. IEEE Micro Top Picks Honorable Mention.

10 Y. Kim et al., "Flipping Bits in Memory without Accessing Them: An Experimental Study of DRAM Disturbance Errors", ISCA 2014. IEEE TCAD Top Pick Paper Award in Hardware Security.

For further information on Professor Mutlu's and the SAFARI Research Group's publications, please see:

https://scholar.google.com/citations?user=7XyGUGkAAAAJ&hl=en https://people.inf.ethz.ch/omutlu/projects.htm

https://safari.ethz.ch/safari-newsletter-january-2021/