**ACM SIGARCH Maurice Wilkes Award goes to Prof. Onur Mutlu**

**Main content**

26.06.2019 | Katharina Emry

**The 2019 ACM SIGARCH Maurice Wilkes Award goes to Prof. Onur Mutlu for "innovative contributions in efficient and secure DRAM systems."**

[Prof.Onur Mutlu](https://people.inf.ethz.ch/omutlu/) is head of the [Institute for High Performance Computing Systems](https://ihpcs.ethz.ch/) at the Department of Computer Science and leads the [SAFARI Research Group](https://safari.ethz.ch/). His research lies in the areas of computer architecture, computer hardware security and energy-efficient computing, with a strong focus on memory systems. This year Prof. Onur Mutlu was honoured with the ACM SIGARCH Maurice Wilkes Award for his “innovative contributions in efficient and secure DRAM systems.”  The notable award is given out each year to a mid-career researcher to recognise "outstanding contributions to computer architecture." The award is named after Maurice V. Wilkes, a British computer scientist and a Turing Award winner, who designed and helped build the Electronic delay storage automatic calculator, one of the earliest stored program computers, and who is the father of microprogramming, which has been used in microprocessors for more than 60 years.

**Seminal research on memory controllers**

After finishing his PhD, he joined Microsoft Research Redmond to start the Computer Architecture Group. During his 2.5-year tenure there, he conducted seminal research into memory systems for multi-core processors. His research into memory controllers spearheaded a fresh research area in computer architecture, which continues to thrive. With his collaborator Thomas Moscibroda, an ETH alumnus, he discovered that existing multi-core memory controllers were vulnerable to denial service attacks [Moscibroda+, USENIX Security’07]. In a series of works published at top venues in computer architecture since 2007, he co-devised new memory control algorithms that provide high system performance, fairness, and quality of service. These techniques turned memory controllers into a center of attention in computer architecture. In particular, his MICRO’07 paper on ”Stall-Time Fair Memory Access Scheduling” and ISCA’08 paper on ”Parallelism-Aware Batch Scheduling” have greatly influenced academic research on memory controllers by exposing new problems and new solutions. Variants of his Parallelism-Aware Batch Scheduler [Mutlu+, ISCA’08] are implemented in some memory controllers designed by Samsung and others.

**Solving the *memory problem* from all angles**

In 2009, he moved to Carnegie Mellon University and later in 2015 to ETH Zurich. He continued attacking the ”memory problem” from all angles, with impact on both major academic research directions and commercial products. He is especially known for his seminal work on DRAM and flash memory, the two major memory technologies used in almost all computing systems today, and emerging memory technologies, such as PCM (phase change memory). Some examples of his work on DRAM technology, for which he is being recognized with the Maurice Wilkes Award, follow.

**The discovery of the RowHammer problem**

In 2014, his group discovered the RowHammer problem [Kim+, ISCA’14], a failure mechanism affecting most real DRAM chips, i.e., memory chips that are used in almost all computing platforms today. This work shook the fundamentals of systems security: RowHammer is the first example of a hardware failure mechanism that causes a practical and widespread system security vulnerability. RowHammer is the phenomenon that repeatedly accessing a row in a modern DRAM chip predictably causes errors in physically-adjacent rows. It is caused by a hardware failure mechanism called read disturb errors, which is a manifestation of circuit-level cell-to-cell interference in a scaled memory technology. Google and others demonstrated attacks that exploit RowHammer to take over an otherwise-completely-secure system. Building on his initial fundamental work that appeared at ISCA 2014, Google Project Zero demonstrated that this hardware phenomenon can be exploited by user-level programs to gain kernel privileges. Many other recent works demonstrated other attacks exploiting RowHammer, including remote takeover of a server vulnerable to RowHammer and takeover of a mobile device by a malicious user-level application that requires no permissions, and many papers continue to be published in top computer science venues on new attacks that exploit the RowHammer vulnerability and new solutions that mitigate it. It continues to have widespread impact on systems, security, software, and hardware communities, both academic and industrial [Mutlu, DATE’17][Mutlu+, IEEE TCAD'19]: for example, it caused a new Hammer Test to be included in standard memory test programs, and Apple cited Onur's work [Kim+, ISCA’14] in its critical security release that introduced a hardware patch to mitigate RowHammer. Various RowHammer solutions he proposed, including Probabilistic Adjacent Row Activation [Kim+, ISCA’14], are implemented in memory controllers and DRAM chips. Due to its widespread impact, Onur's RowHammer work was recently recognized as one of the seven papers of 2012-2017 selected as Top Picks in Hardware and Embedded Security. More information on RowHammer can be found in his recent retrospective paper entitled [**"RowHammer: A Retrospective"**](https://arxiv.org/pdf/1904.09724.pdf).

**More efficient, more reliable, faster memory**

Onur is one of the pioneers of architectural research on solving critical DRAM scaling problems, enabling memory chips to become higher performance, more efficient, and more reliable. He experimentally demonstrated, analyzed, and provided architectural solutions for critical issues (refresh, latency, variability) by analyzing modern DRAM chips using real FPGA-based experimental platforms. He showed that refresh is a major scaling challenge and performance/energy limiter of future DRAM chips [ISCA '12]. He demonstrated key problems affecting the identification of data retention times in modern DRAM chips and showed difficulties in practically implementing various architectural solutions [ISCA '13], providing device-level data available nowhere else. He developed new online profiling mechanisms to solve the problem by adaptively determining refresh rates of different DRAM rows [SIGMETRICS '14, DSN '15, ISCA/MICRO '17], effectively saving energy and improving performance at the same time. He experimentally showed that DRAM latencies can be significantly reduced by adapting access latency to common-case operating conditions [HPCA '15] and latency characteristics of different memory parts [SIGMETRICS '16, '17]. Intel and Samsung advocated [Memory-Forum '14] that his subarray-level parallelism idea, enabling a large reduction in latency of conflicting memory requests [ISCA '12], be part of future DRAM standards to tolerate increasing DRAM latencies caused by technology scaling.

**Security primitives for memory and hardware**

His work experimentally demonstrated, analyzed, and provided architectural solutions for various other critical DRAM issues (e.g., latency [Lee+, HPCA’13, HPCA’15, SIGMETRICS’17][Hassan, HPCA'16][Hassan, ISCA'19], variability [Chang+, SIGMETRICS’16][Kim+, HPCA’18], energy [David+, ICAC’11][Chang+, SIGMETRICS’17], power [Ghose+, SIGMETRICS’18], reliability [Meza+, DSN’15]) by analyzing modern DRAM chips using real FPGA-based experimental platforms [Hasan+, HPCA’17], providing precious data available nowhere else. His group has most recently shown that DRAM memory can be enabled as a substrate to support various important security primitives: to generate true random numbers at high speed and low energy [HPCA'19], to enable physically unclonable functions [HPCA'18], which are important for system authentication, and to quickly destroy in-memory data when under attack. He continues to research hardware security primitives and solutions that can make memories and thus entire computing platforms more secure.

**A new computing paradigm**

Onur has been researching how to build fundamentally more efficient systems by changing the computing paradigm, such that memory devices can compute. For the entire computing history, memory devices have remained as dumb data storage units that cannot compute, leading to a huge data movement bottleneck between the processor and memory that plagues all computing systems, causing great inefficiency, energy waste, and performance loss. Onur's efficient DRAM research aims to change that completely by enabling DRAM devices to accelerate key computations internally. His work showed, for the first time in literature, that DRAM devices can be enabled to perform fundamental operations such as copy, initialization [MICRO'13], and bitwise AND, OR, NOT [MICRO'17]. Doing so greatly improves both latency and energy efficiency of such fundamental operations, by 1-2 orders of magnitude. His Ambit substrate [MICRO'17] is the first to be able to execute *any* application using DRAM chips at low cost. His results show that database query latencies can be reduced by an order of magnitude with this substrate. Onur's recent work also shows that other forms of processing in memory with the emerging 3D-stacked technology can greatly improve both performance and energy of many workloads in many different computing platforms [ISCA'15a,b, ASPLOS'18, ISCA'19], sometimes by more than an order of magnitude. A large focus of his research group at ETH Zurich is currently on enabling such a fundamental paradigm shift in the way computers are designed, by processing data close to where the data resides, i.e., in memory or storage. A recent work he has authored with his senior researchers summarizes the goals, benefits, and challenges of such an approach and can be found [here](https://people.inf.ethz.ch/omutlu/pub/ProcessingDataWhereItMakesSense_micpro19-invited.pdf) as an invited paper entitled "Processing Data Where It Makes Sense: Enabling In-Memory Computation."