AN IN-MEMORY COMPUTING SERIES

Next Talk: 3/May/2021, 4-5:30pm CET

INTELLIGENT ARCHITECTURES FOR INTELLIGENT MACHINES

Professor Onur Mutlu, ETH Zurich

Computing is bottlenecked by data. Large amounts of application data overwhelm storage capability, communication capability, and computation capability of the modern machines we design today. As a result, many key applications' performance, efficiency and scalability are bottlenecked by data movement. We describe three major shortcomings of modern architectures in terms of 1) dealing with data, 2) taking advantage of the vast amounts of data, and 3) exploiting different semantic properties of application data. We argue that an intelligent architecture should be designed to handle data well. We show that handling data well requires designing architectures based on three key principles: 1) data-centric, 2) data-driven, 3) data-aware. We give several examples for how to exploit each of these principles to design a much more efficient and high performance computing system. We will especially discuss recent research that aims to fundamentally reduce memory latency and energy, and practically enable computation close to data, with at least two promising novel directions: 1) performing massively-parallel bulk operations in memory by exploiting the analog operational properties of memory, with low-cost changes, 2) exploiting the logic layer in 3D-stacked memory technology in various ways to accelerate important data-intensive applications. We discuss how to enable adoption of such fundamentally more intelligent architectures, which we believe are key to efficiency, performance, and sustainability. We conclude with some guiding principles for future computing architecture and system designs.

More information about the event and the speaker: https://www.ict.tuwien.ac.at/staff/taherinejad/MiM/next.html Mondays in Memory (MIM) is a free biweekly webinar series open to everyone around the world and dedicated to all aspects and technologies related to in-memory computing (including, in a broader sense, near-memory computing too). MIM will be held on the first and third Monday of each month (starting in May 2021) at 4pm CET (7am Pacific time, and 10pm Beijing time).

Each webinar starts with a 40mins talk by a speaker, followed up with a 40mins questions and discussions with the speaker and two panel members. Dr. Nima Taherinejad hosts the webinars, and together with his team they organize the MiM series.

Website: http://www.ict.tuwien.ac.at/ staff/ aherinejad/MiM/ Email:nima.taherinejad@tuwien.ac.at

Onur Mutlu is a Professor of Computer Science at ETH Zurich. He is also a faculty member at Carnegie Mellon University,



where he previously held the Strecker Early Career Professorship. His current broader research interests are in architecture, systems, computer hardware security, and bioinformatics. A variety of techniques he, along with his group and collaborators, has invented over the years have influenced industry and have been employed commercial microprocessors and memory/storage systems. He obtained his PhD and MS in ECE from the University of Texas at Austin and BS degrees in Computer Engineering and Psychology from the University of Michigan, Ann Arbor. He started the Computer Architecture Group at Microsoft Research (2006-2009), and held various product and research positions at Intel Corporation, Advanced Micro Devices, VMware, and Google. He has received numerous awards and recognitions. For more information, please see his page https://people.inf.ethz.ch/omutlu/ his group page https://safari.ethz.ch/ and his Youtube channel at https://www.youtube.com/ **OnurMutluLectures**