

Computer Science Seminar Series

National Capital Region

Dynamically Heterogeneous Cores through 3D Resource Pooling

Speaker: Prof. Houman Hodayoun
George Mason University

Friday, February 7, 2014

1:00PM- 2:00PM, NVC 325

Abstract

3D die stacking is a recent technological development which makes it possible to create chip multiprocessors using multiple layers of active silicon bonded with low latency, high-bandwidth, and very dense vertical interconnects. 3D die stacking technology provides very fast communication, as low as a few picoseconds, between processing elements residing on different layers of the chip. The rapid communication network in a 3D stack design, along with the expanded geometry, provides an opportunity to dynamically share on-chip resources among different cores. This research describes an architecture for a dynamically heterogeneous processor architecture leveraging 3D stacking technology. Unlike prior work in the 2D plane, the extra dimension makes it possible to share resources at a fine granularity between vertically stacked cores. As a result, each core can grow or shrink resources, as needed by the code running on the core. This architecture, therefore, enables runtime customization of cores at a fine granularity and enables efficient execution at both high and low levels of thread parallelism. This architecture achieves performance gains of up to 2X, depending on the number of executing threads, and gains significant advantage in energy efficiency.

Biography



Houman Hodayoun is an Assistant Professor of the Department of Electrical and Computer Engineering at George Mason University. He also holds a Courtesy appointment with the Department of Computer Science. Prior to joining George Mason University, He spent two years at the University of California, San Diego, as National Science Foundation Computing Innovation (CI) Fellow awarded by the Computing Research Association (CRA) and the Computing Community Consortium (CCC). Houman research is on power-temperature and reliability-aware memory and processor design optimizations and spans the areas of computer architecture, embedded systems, circuit design, and VLSI-CAD, where he has published more than 40 technical papers on the subject, including some of the earliest work in the field to address the importance of cross-layer power and

temperature optimization in memory peripheral circuits. He is currently leading a number of research projects, including the design of next generation 3D heterogeneous multicores, low power hybrid SRAM-NVM memory hierarchy design, reliability-aware cache design, and power management in data centers. Houman was a recipient of the four-year University of California, Irvine Computer Science Department chair fellowship. He received his PhD degree from the Department of Computer Science at the University of California, Irvine in 2010, an MS degree in computer engineering in 2005 from University of Victoria, Canada and his BS degree in electrical engineering in 2003 from Sharif University of technology.