Seminar – Tuesday, July 18, in 111 Rhodes Annex **2:00 – 3:00 p.m.**

Holcombe Department of Electrical and Computer Engineering Seminar Series

Viewing the Computer as a Software-Defined Network: Programmable Architectures for Resourcing-on-Demand

Dr. Sally McKee

Professor, Chalmers University of Technology in Gothenburg, Sweden

Abstract

This talk describes a novel approach to Computer Architecture: viewing the computer as a Software-Defined Network that can provide differentiated services for different applications. The key insight is that modern processors are mostly wires, and the NoC that connects the growing number of cores and accelerators already uses packets for the communications among components. To deal with the semantic gap between high-level applications and underlying hardware packets, we can attach a high-level semantic tag (e.g., a virtual machine or thread ID) to each memory-access, I/O, or interrupt packet. We can then implement a programmable control plane that can be integrated into shared architectural resources (e.g., caches, DRAM, and I/O devices) and can process packets differently according to tag-based rules. By splitting the NoC into a control plane and a data plane, we enable things like:
+ differentiated services/more efficient QoS (there'll be less need for over-provisioning),
+ real-time monitoring of performance/power/faults,
+ more efficient power-aware runtimes,
+ efficient tracking of information flow for security,
+ hardware-supported virtualization, and
+ selective approximation at all levels of the architecture.

The talk addresses how this new perspective relates to my past work and how it shapes the research program I plan to build, particularly with respect to potential collaborations.

Biography of Speaker

McKee received her bachelor's degree from Yale, master's from Princeton, and doctorate from the University of Virginia. Her dissertation advisor was Bill Wulf, with whom she worked on memory systems architecture. Together they coined the now-common term "the memory wall" to describe a situation in which processors are always waiting on memory, and CPU performance is therefore entirely limited by memory performance.

Before graduate school, McKee worked for Digital Equipment Corporation and Microsoft Corporation. She also held internships at Digital Equipment Corporation's Systems Research Center (now HP Labs) and the former AT&T Bell Labs. After her Ph.D. and postdoc, McKee joined Intel's Microcomputer Research Lab. During her time at Intel, she taught at the Oregon Graduate Institute and Reed College. From 1998 to 2002, she was a Research Assistant Professor at the University of Utah's School of Computing, where she worked on the Impulse Adaptable Memory Controller project. She then joined the Computer Systems Lab within Cornell's School of Electrical and Computer Engineering. Since November, 2008, McKee has been with Chalmers University of Technology in Gothenburg, Sweden.

Her research has historically focused on analyzing application memory behavior and designing more efficient memory systems together with the software to exploit them. Achieving this broad objective requires developing new underpinnings for system understanding, and thus she and her students and collaborators have: developed new approaches to performance analysis; built scalable tools for application analysis and system modeling; designed architectures to enable more comprehensive system introspection and analyses; designed efficient memory systems for both parallel and uniprocessor (including embedded) platforms; and automated memory optimizations for HPC applications. More recent work does many of the same things for data centers and tries to better understand where energy is spent in modern servers.