



Rethinking the Architecture of Warehouse-Scale Computers: Improving Efficiency and Utilization

Jason Mars

University of California in San Diego

jmars@eng.ucsd.edu

The class of datacenters coined as “warehouse scale computers” (WSCs) house large-scale data intensive web services such as websearch, maps, social networking, docs, video sharing, etc. Companies like Google, Microsoft, Yahoo, and Amazon spend ten to hundreds of millions to construct and operate WSCs to provide these services. Maximizing the efficiency of this class of computing reduces cost and has energy implications for a greener planet. However, WSC design and architecture remains in its relative infancy.

WSCs are built using commodity processor architectures (Intel/AMD), and software components (Linux, GCC, JVM, etc) that has been engineered and optimized for traditional computing environments and workloads, such as those you’d find in the desktop/laptop environment. However, there are many characteristics, assumptions, and requirements present in the WSC computing domain that impacts design decisions within these components. In this presentation, we rethink how WSCs are designed and architected, identify sources of inefficiency, and develop solutions to improve WSCs, with a particular focus on the interaction between the application layer, system software stack, and the underlying hardware platform.

Biography

Jason Mars is currently an Assistant Professor at the University of California in San Diego. He received his Ph.D. of Computer Science at the University of Virginia in 2012. He has been an active researcher in the areas of computer architecture, system software, and cross-layer system design. Jason has published 24 papers in these areas and received a number of rewards and honors for excellence in his research work. You can find out more information about him at <http://www.jasonmars.org/>