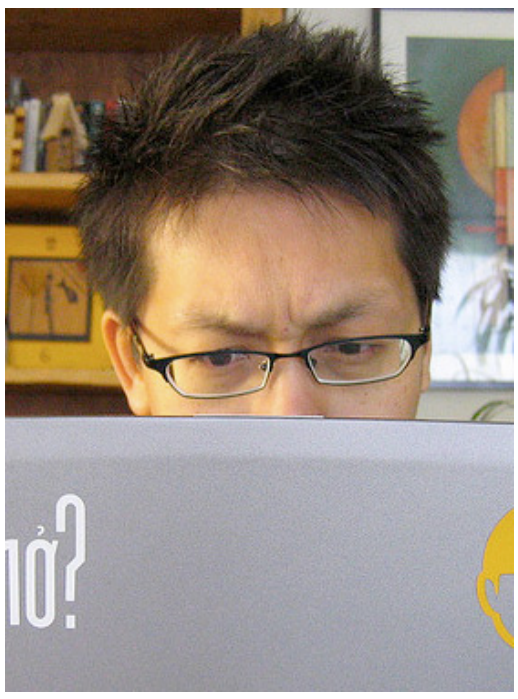
Electrical & Computer ENGINEERING **SEMINAR**

Thursday, October 27

Scaife Hall Auditorium
Room 125 at 4:30 p.m.
Refreshments at 4:00 p.m.



Richard Vuduc
Assistant Professor
School of Computational Science and Engineering
Georgia Tech

Richard (Rich) Vuduc is an assistant professor in the School of Computational Science and Engineering at Georgia Tech. His research lab, the HPC Garage (hpcgarage.org), is interested in high-performance computing with focus areas in parallel algorithms, performance analysis, tuning, and debugging. He received an NSF CAREER award in 2010 and was an invited member of DARPA's 2009-2010 Computer Science Study Group. Most recently, he was a member of the Georgia Tech team that won the 2010 Gordon Bell Prize.

Balance principles for algorithm-architecture co-design

Given an algorithm, what processor and memory architecture will deliver the best performance and power/energy efficiency? Conversely, given an architecture, what class of algorithms will (or will not) run efficiently? We are investigating analytical frameworks to answer these kinds of co-design questions, based on the concept of balance principles. A balance principle is a theoretical constraint equation that explicitly relates algorithm parameters to hardware parameters according to some figure of merit, such as speed, power, or cost. This notion originates in early work by Kung (1986) and numerous others; we reinterpret the classical notions of balance in a modern context of parallel and I/O-efficient algorithm design as well as trends in emerging architectures. From such a principle, we argue that one can better understand algorithm and hardware trends, and furthermore gain insight into how to improve both algorithms and hardware.

We use this principle to make quantitative predictions about future supercomputer systems as we proceed toward exascale machines. These predictions include when matrix multiply might become memory-bound; for what algorithms stacking processors and memory will be beneficial,and for which it will not; and whether supercomputers based on embedded CPU-like processors or those based on GPU-like processors will be better and why. Our overall aim is to suggest how to one might co-design rigorously and quantitatively while still yielding intuition and insight.

ECE Seminar Hosts

Gabriela Hug	ghug@ece.cmu.edu
Lujo Bauer	lbauer@cmu.edu
Soummya Kar	soummyak@andrew.cmu.edu
Jeff Weldon	jweldon@ece.cmu.edu

Carnegie Mellon

Electrical & Computer ENGINEERING