The University of Alabama
Department of Computer Science
Colloquium Series

Dr. Philip Rhodes

University of Mississippi
Department of Computer and Information Science

Spatial Data, Latency, Parallel Computing and Visualization

Monday, October 12th
11:00 a.m., East Engineering Room 110

Abstract:
Large spatial datasets present special challenges for efficient access and representation. A familiar example of a spatial dataset is the MRI and CT data that comprises a volumetric representation of a patient's body. However, gains in computational power have made possible the simulation of spatial phenomena in a wide variety of fields such as climatology, physics, engineering, and seismology. The importance of efficient access to spatial data will only grow as the popularity and size of such datasets increases.

While disk and network bandwidth have steadily improved over recent decades, *latency* continues to hinder performance. This problem is especially difficult with spatial datasets because elements that are nearby in the dataspace may be stored far apart on disk, requiring multiple read operations, incurring multiple disk latency penalties. In distributed computing environments, network latency penalties can be an order of magnitude larger than disk latency, and their number must be minimized to achieve acceptable performance.

Iteration Aware Prefetching (IAP) addresses this problem by using advance knowledge of an access pattern to transform a large number of small transactions into a small number of large transactions. Since the total number of transactions is reduced, fewer latency penalties are paid, and performance is improved significantly. Part of this talk will discuss the use of different forms of IAP in data visualization and other applications.

It is natural to employ parallel computing when dealing with large datasets, due to the amount of processing involved. Cluster computing is motivated by the cost effectiveness of parallel computing with commodity machines. More recently, Graphics Processing Units (GPUs) have evolved into surprisingly powerful parallel computers, containing tens to even hundreds of cores in recent models. For both cluster and GPU computing, data access must be done efficiently in order to keep multiple processors busy. A second topic of this talk will be the development of *IDEA,* an API for parallel programming with spatial data that frees the scientist/programmer from the minutiae of efficient I/O for cluster computing. If time permits, we will also look at GPU computing, and the importance of addressing latency in that exciting new programming natural world.

Bio:
Dr. Rhodes is an Assistant Professor in the Department of Computer and Information Science at the University of Mississippi. Funded by the National Science Foundation and the Department of Homeland Security, his research focuses on parallel computing, efficient access to spatial datasets in distributed and parallel computing environments, and also includes related issues in visualization and computer graphics.

Dr. Rhodes received his Ph.D. in Computer Science from the University of New Hampshire in 2004. He also holds degrees from the University of Rhode Island and the University of Virginia.