Notice and Invitation
Oral Defense of Doctoral Dissertation
The Volgenau School of Engineering, George Mason University

Maria Malik
Bachelor of Science, Center for Advanced Studies in Engineering, 2008
Master of Science, George Washington University, 2011

System, Architectural and Application level analysis of Big Data Applications for Performance and Energy-Efficiency
Tuesday, March 20, 2018, 11:00 am-12:00 pm
Engineering Building, Room 3507
All are invited to attend.

Committee
Dr. Houman Homayoun
Dr. Kris Gaj
Dr. Jens Peter Kaps
Dr. Setareh Rafatirad

Abstract

The world of big data is changing constantly that creates challenges to process the applications using existing solutions. Big data applications require computing resources and storage subsystems that can scale to manage massive amounts of diverse data. Furthermore, physical design constraints have become the dominant limiting factor for scaling out servers. Therefore recent work advocates the use of low-power embedded cores in servers to address these challenges.

In this research, first through comprehensive system and micro-architectural level analysis, we characterize the big data applications on big Xeon and little Atom-based server architecture to demonstrate how the choice of big vs little core-based server for energy-efficiency is influenced by the size of data, performance constraints, and presence of accelerator. Second, through methodical investigation of performance and power measurements, we demonstrate how the interplay among various MapReduce configurations as well as application and architecture level parameters create new opportunities to co-locate MapReduce applications at the node level.

Co-locating and self-tuning of unknown applications are challenging problems, especially when co-locating multiple big data applications concurrently with many tuning knobs, potentially requires exhaustive brute-force search to find the right settings. This research challenge upsurges an imminent need to develop a technique that co-locates applications at a node level and predict the optimal system, architecture and application level configure parameters to achieve the maximum energy efficiency. Towards this goal, we develop an Energy-Efficient Co-Locating and Self-Tuning (ECoST) technique for co-located MapReduce applications to enhance their energy-efficiency.