Emerging Opportunities, Challenges, and Applications in Exascale Computing

Ananth Grama
Professor of Computer Science
Purdue University
Friday, February 18, 2011
3:00 pm, Birck 1001

Abstract: The move towards exascale computing platforms (capable of $10^{18}$ floating point operations) poses tremendous challenges, while presenting opportunities for foundational advances in a variety of application domains. In this talk, I will describe various technical issues that must be overcome to realize the potential of exascale computing platforms. It is widely believed that hardware underlying these platforms will have upwards of $10^7$ computing cores, is likely to integrate SIMD (GPU)/FPGA coprocessors, have high latency interconnects (w.r.t. CPU clock), and is unlikely to have significantly higher RAM/core than today's HPC platforms. Issues of energy efficiency (watts/GFLOP) and fault tolerance/relience (MTBF at high component count) are prime. This is in addition to more traditional challenges associated with memory/storage bandwidth and capacity, latency tolerance, and concurrency at the extreme, which are all exacerbated by orders of magnitude. Exascale computing algorithms may differ significantly from serial counterparts in being suboptimal in terms of operation counts, but be tolerant to high latencies and relaxed synchronization. Proving stability and performance of these algorithms at scale is likely to emerge as an issue. Implementations of these algorithms on heterogeneous platforms will require a hierarchical design paradigm -- leveraging intra-node and inter-node concurrency through potentially different programming models. Finally, the success of these platforms will be determined by the new science and applications it enables. I will conclude with a brief discussion of various consortia and organizations focused on exascale computing related activities.

Bio: Professor Grama's research interests span the areas of parallel and distributed computing architectures, algorithms, and applications. His work on distributed infrastructure deals with development of software support for dynamic clustered and multiclaned environments. More recent work has focused on resource location and allocation mechanisms in peer-to-peer networks. His research on applications has focused on particle dynamics methods, their applications to dense linear system solvers, and fast algorithms for data compression and analysis.

Professor Grama has authored several papers and co-authored a text book Introduction to Parallel Computing: Design and Analysis of Algorithms with Vipin Kumar, Anshul Gupta, and George Karypis. He is a member of American Association for Advancement of Sciences and Sigma Xi.