ABSTRACT OF THE DISSERTATION

The Interplay Between Energy Efficiency and Resilience for Scalable High Performance
Computing Systems

by

Li Tan

Doctor of Philosophy, Graduate Program in Computer Science
University of California, Riverside, December 2015
Dr. Zizhong Chen, Chairperson

As the exascale supercomputers are expected to embark around 2020, supercomputers nowadays expand rapidly in size and duration in use, which brings demanding requirements of energy efficiency and resilience. These requirements are becoming prevalent and challenging, considering the crucial facts that: (a) The costs of powering a supercomputer grow greatly together with its expanding scale, and (b) failure rates of large-scale High Performance Computing (HPC) systems are dramatically shortened due to a large amount of compute nodes interconnected as a whole. It is thus desirable to consider both crucial dimensions for building scalable, cost-efficient, and robust HPC systems in this era. Specifically, our goal is to fulfill the optimal performance-power-failure ratio while exploiting parallelism during HPC runs.

Within a wide range of HPC applications, numerical linear algebra matrix operations including matrix multiplication, Cholesky, LU, and QR factorizations are fundamental and have been extensively used for science and engineering fields. For some scientific applications, these matrix operations are the core component and dominate the total execution

time. Saving energy for the matrix operations thus significantly contributes to the energy efficiency of scientific computing nowadays. Typically, when processors are experiencing idle time during HPC runs, i.e., slack, energy savings can be achieved by leveraging techniques to appropriately scale down processor frequency and voltage during underused execution phases. Although with high generality, existing OS level energy efficient solutions can effectively save energy for some applications in a black-box fashion, they are however defective for applications with variable workloads such as the matrix operations – the optimal energy savings cannot be achieved due to potentially inaccurate and high-cost workload prediction they rely on. Therefore, we propose to utilize algorithmic characteristics of the matrix operations to maximize potential energy savings. Specifically, we achieve the maximum of energy savings in two ways: (a) reducing the overhead of processor frequency switches during the slack, and (b) accurately predicting slack of processors via algorithm-based slack prediction, and eliminating the slack accordingly by respecting the critical path of an HPC run.

While energy efficiency and resilience issues have been extensively studied individually, little has been done to understand the interplay between them for HPC systems. We propose to quantitatively analyze the trade-offs between energy efficiency and resilience in the large-scale HPC environment. Firstly, we observe that existing energy saving solutions via slack reclamation are essentially frequency-directed, and thus fail to fully exploit more energy saving opportunities. In our approach, we decrease the supply voltage associated with a given operating frequency for processors to further reduce power consumption at the cost of increased failure rates. We leverage the mainstream resilience techniques to

tolerate the increased failures caused by the undervolting technique. Our strategy is theoretically validated and empirically evaluated to save more energy than a state-of-the-art frequency-directed energy saving solution, with the guarantee of correctness. Secondly, for capturing the impacts of frequency-directed solutions and undervolting, we also develop analytic models that investigate the trade-offs among resilience, energy efficiency, and scalability for large-scale HPC systems. We discuss various HPC parameters that inherently affect each other, and also determine the optimal energy savings at scale, in terms of the number of floating-point operations per Watt, in the presence of undervolting and fault tolerance.