

Zusammenfassung

Energieeffizientes Rechnen wurde als eine der zehn größten Herausforderungen in der Exascale-Forschung identifiziert. Um eine hohe Energieeffizienz zu erreichen, werden Fortschritte bei der Hardware allein nicht ausreichen. Erhebliche Fortschritte bei den Softwaretechnologien sind ebenfalls erforderlich, um sicherzustellen, dass Exascale-Systeme eine hohe Leistung bei effektivem Energiemanagement erreichen. Eine verbreitete Technik ist die dynamische Spannungs- und Taktfrequenzskalierung (Dynamic Voltage and Frequency Scaling, DVFS), eine Energiesparstrategie, die die Leistungs- und Taktungseinstellungen eines Prozessors anpasst. Die Implementierung von DVFS ist jedoch aufgrund der großen Anzahl von Taktfrequenzen, der unterschiedlichen Energiecharakterisierung verschiedener Anwendungen, der Komplexität von Multi-Kernel-Applikationen und der fehlenden Portabilität zwischen verschiedenen Herstellern nicht einfach.

In dieser Arbeit untersuchen wir die zuvor genannten Herausforderungen bei GPUs und schlagen fortschrittliche DVFS-Techniken vor, um Engpässe zu beseitigen und das energieeffiziente Rechnen zu verbessern. In unserer Forschung werden neue Energiemodellierungsansätze auf der Grundlage des maschinellen Lernens zur Vorhersage des Energieverbrauchs vorgeschlagen, ausgehend von einem einzelnen Kernel und später für Multitasking-Anwendungen. Wir schlagen außerdem SYnergy vor, ein neuartiges energieeffizientes Framework, das Sprachen, Compiler, Laufzeiten und Job Scheduler umfasst, um beispiellose, feinkörnige Energieeinsparungen auf großen heterogenen Clustern zu erzielen. Insgesamt haben die Ergebnisse dieser Arbeit zu vier Veröffentlichungen beigetragen, darunter der Top-Konferenz: ICPP und SC. Die Ergebnisse der Dissertation wurden auch in zwei industriellen Anwendungen im Rahmen des EuroHPC-Projekts umgesetzt, insbesondere in der Ligen Drug Discovery Pipeline (LIGATE) und in Cronos.

Abstract

Energy-efficient computing has been identified as one of the top ten Exascale research challenges. Toward achieving high energy efficiency, hardware advances will not be sufficient in isolation. Significant advances in software technologies are also required to guarantee that Exascale systems achieve high performance with effective energy management. One popular technique is Dynamic Voltage and Frequency Scaling (DVFS), an energy-saving strategy that adjusts power and frequency settings on a processor. However, DVFS implementation is not easy because of the large number of frequency settings, the different energy characterization of different applications, the complexity of multi-kernel applications, and the non-portability across different hardware vendors.

In this thesis, we investigate the above challenges in GPUs and propose DVFS-aware advanced techniques that tackle these bottlenecks. Our research proposes new energy modeling approaches based on machine learning to predict energy consumption, starting from a single kernel and later attacking multi-task applications. We also propose SYnergy, a novel energy-efficient framework that spans languages, compilers, runtimes, and job schedulers to achieve unprecedented fine-grained energy savings on large-scale heterogeneous clusters. Overall, the results of this thesis have contributed to four publications, including the top conference: ICPP and SC. The thesis results have also been implemented on two industrial applications within the EuroHPC project, specifically, on the Ligen drug discovery pipeline (LIGATE) and on Cronos.