



MASTERARBEIT

„Leistungsmodellierung im HPC“

Stefan de Souza

Elitestudiengang „Scientific Computing“

Universität Bayreuth, 2024

Energie und Leistung auf HPC-Systemen

Stefan de Souza hat im Elitestudiengang "Scientific Computing" an der Universität Bayreuth studiert. Im Rahmen einer Kooperation mit dem RRZE (Regionales Rechenzentrum Erlangen) hat er seine Masterarbeit zur Leistungs- und Energiemodellierung auf modernen Multicore-Server-CPU's auf den dortigen Rechenclustern bearbeitet.

Beweggründe für die Studie

Im Zeitalter des Exascale Computing, in dem ein einziger CPU-Chip Hunderte von physischen Kernen beherbergen kann, ist der Leistungs- und Energieverbrauch von HPC-Systemen zu einem wichtigen Thema geworden. Vor allem in Anbetracht der steigenden Energiekosten in Europa ist die Optimierung zur Erzielung der größtmöglichen Leistung pro Euro jetzt von noch größerem Interesse. Dies erfordert ein Verständnis der vielschichtigen Natur des Stromverbrauchs und wie die Codeleistung damit zusammenhängt. Die meisten wissenschaftlichen Codes lassen sich in kleinere Schleifenkerne aufteilen, in denen sowohl Berechnungen als auch Kommunikation durchgeführt werden. Daher wäre es äußerst nützlich, die Verlustleistung dieser kleineren Codeteile im Voraus vorhersagen zu können, ohne sie tatsächlich ausführen zu müssen.

Wissenschaftliche Modellierung

Beim Versuch einer wissenschaftlichen Modellierung der Verlustleistung gibt es im Allgemeinen zwei grundlegende Ansätze. Die erste Möglichkeit beinhaltet Annahmen über die zugrundeliegenden physikalischen Eigenschaften (z. B. das Verlustverhalten der beeinträchtigten Schaltkreise), was zu einem "Bottom-up"-Ansatz führt. Dagegen beruht der zweite Ansatz auf gemessenen Metriken von tatsächlichen Leistungszählern der CPU, um wichtige Konfigurationsfaktoren zu ermitteln. Dies erfordert wesentlich weniger Wissen und funktioniert als "Top-Down"-Ansatz.

Erzielung eines genauen Modells

Angesichts der Fülle der verfügbaren wissenschaftlichen Benchmarks können auch statistische Methoden eingesetzt werden, indem Regressionen mit Datensätzen über verschiedene Systemkonfigurationen hinweg durchgeführt werden. Diese Benchmarks wurden auf Assembler-Kernen implementiert, um verfrühte Optimierungen zu vermeiden, die bei Compilern in höheren Programmiersprachen häufig auftreten. Letztlich ermöglicht ein Mittelweg zwischen der Berücksichtigung von bekanntem Verhalten, Mikrobenchmarks und angepassten Parametern die Abschätzung der Verlustleistung für unbekannte Codeteile innerhalb einer kleinen Fehlertoleranz auf einem bestimmten CPU-Knoten.

Mehr zum Elitestudiengang:

<https://www.elitenetzwerk.bayern.de/start/foerderangebote/elitestudiengaenge/uebersicht-elitestudiengaenge/scientific-computing>

<https://www.scientific-computing.uni-bayreuth.de/en/index.html>