

Research Project

MLS: Multilevel Scheduling in Large Scale High Performance Computers (extension)

Project funded by own resources

Project title MLS: Multilevel Scheduling in Large Scale High Performance Computers (extension)

Principal Investigator(s) Ciorba, Florina M. ;

Project Members Muller Korndorfer, Jonas Henrique ;

Organisation / Research unit

Departement Mathematik und Informatik / High Performance Computing (Ciorba)

Project Website <https://hpc.dmi.unibas.ch/en/research/mls/>

Project start 01.09.2020

Probable end 30.04.2021

Status Completed

This project proposes to investigate and develop multilevel scheduling (MLS), a multilevel approach for achieving scalable scheduling in large scale high performance computing systems across the multiple levels of parallelism, with a focus on software parallelism. By integrating multiple levels of parallelism, MLS differs from hierarchical scheduling, traditionally employed to achieve scalability within a single level of parallelism. MLS is based on extending and bridging the most successful (batch, application, and thread) scheduling models beyond single or a couple of parallelism levels (scaling across) and beyond their current scale (scaling out).

The proposed MLS approach aims to leverage all available parallelism and address hardware heterogeneity in large scale high performance computers such that execution times are reduced, performance targets are achieved, and acceptable efficiency is maintained. The methodology for reaching the multi-level scheduling aims involves theoretical research studies, simulation, and experiments.

The expected outcome is an answer to the following research question: Given massive parallelism, at multiple levels, and of diverse forms and granularities, how can it be exposed, expressed, and exploited such that execution times are reduced, performance targets (e.g., robustness against perturbations) are achieved, and acceptable efficiency (e.g., tradeoff between maximizing parallelism and minimizing cost) is maintained? This proposal leverages the most efficient existing scheduling solutions to extend them beyond one or two levels, respectively, and to scale them out within single levels of parallelism. The proposal addresses four tightly coupled problems: scalable scheduling, adaptive and dynamic scheduling, heterogeneous scheduling, and bridging schedulers designed for competitive execution (e.g., batch and operating system schedulers) with those for cooperative execution (e.g., application level schedulers).

Overall, the project aims to make a fundamental advance toward simpler to use large scale high performance computing systems, with impacts not only in the computer science community but also in all computational science domains.

Keywords parallel and distributed computing; large scale computing; multilevel parallelism; concurrency and data locality; dynamic load balancing; scheduling and load balancing; single-level scheduling; high performance computing; multilevel scheduling

Financed by

Other funds

Follow-up Project of... [3623894 MLS: Multilevel Scheduling in Large Scale High Performance Com-](#)

puters

Add publication

Add documents

Specify cooperation partners