

Research Project

Multi-level Scheduling in High Performance Computing (MLS in HPC)

Project funded by own resources

Project title Multi-level Scheduling in High Performance Computing (MLS in HPC)

Principal Investigator(s) Ciorba, Florina M. ;

Project Members Eleliemy, Ahmed Hamdy Mohamed ;

Organisation / Research unit

Departement Mathematik und Informatik / High Performance Computing (Ciorba)

Project start 01.08.2015

Probable end 31.07.2020

Status Completed

The need for large-scale computations or for supporting large data-intensive calculations leads to the use of multiple (clusters of) parallel computers at different sites distributed across the Internet. The computational grid and cloud are examples of such distributed (possibly heterogeneous) computing systems and offer multiple hierarchical levels of parallelism: site, cluster, node, socket, core, vector, pipeline, and instruction [1]. Each level of parallelism requires at least a scheduler. For instance, at the cluster level there are batch schedulers and runtime systems. Depending on the level of parallelism, schedulers can be viewed as global and local. From a site level parallelism perspective, global schedulers distribute the computational tasks or the communication among the different sites, whereas local schedulers distribute the tasks or the communication among the computational nodes of a particular site. Furthermore, from the cluster level parallelism perspective, decisions made by the runtime system regarding the initial placement of application tasks to locally assigned computing resources can significantly influence the outcome of a cluster level scheduler. The scheduling goals differ from level to level and may be conflicting between levels. For instance, cluster level schedulers typically aim at maximizing fairness among all applications in terms of their execution time which may result in non-optimal execution times for certain applications. Application level schedulers typically aim at minimizing the execution time of a single application, which may result in non-balanced execution times among applications. Addressing the problem jointly at multiple levels is called multi-level scheduling and constitutes a multi-objective combinatorial optimization problem.

[1] Michael Wolfe, Compilers and More: Programming at Exascale, March 2011, http://www.hpcwire.com/2011/03/08/com-at_exascale/

Keywords multi-level scheduling, multi-scale scheduling, performance, energy

Financed by

University funds

Add publication

Published results

3716004, Eleliemy, Ahmed; Mohammed, Ali; Ciorba, Florina M., Simulating Batch and Application Level Scheduling Using GridSim and SimGrid, Publication: Other Publications (Forschungsberichte o. ä.)

Add documents

Specify cooperation partners