Laboratory for Parallel Programming

TECHNISCHE UNIVERSITÄT DARMSTADT

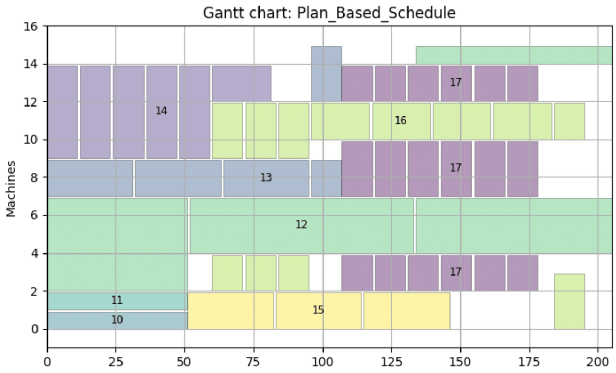
Master's Thesis

<Plan-Based Scheduling Algorithms for HPC Applications>

Motivation

The scheduling of parallel applications has naturally emerged driven by the recent advances in the development of larger HPC clusters. In that direction, and given both the user and system requirements, it is critical to develop an HPC scheduler having in mind a particular set of Quality of Services (QoSs) such as resource utilization and application response time. Furthermore, the energy efficiency of the schedule is a problem of paramount importance to mitigate any negative environmental impact. Such stringed QoSs are known to be contradicting, hence, a trad-off solution is required.

The current state-of-the-art reveals a strong immatureness and an evident lack of solutions concerning all the above-mentioned stringent QoSs. Such immatureness is mainly due to the uncertainty in the user application requirements such as CPU, Memory, and IO requirements. Therefore, this topic aims essentially at further developing the support for HPC schedulers by designing and implementing resilient optimization scheduling algorithms, known as plan-based scheduling algorithms, so that the desired QoSs are met.



Plan-Based schedule for HPC applications.

Task

Plan-based scheduling algorithms create an execution plan of all pending applications submitted by the users. Hence, plan-based schedulers rely on optimization algorithms to return optimal/near-optimal schedule plan that improves a particular set of QoS. To successfully finish the work, the following tasks should be completed:

- Study the existing literature related to the design of HPC job schedulers.
- Mathematical formation of the scheduling problem.
- Design/implementation of optimization scheduling algorithms that return optimal/sub optimal solutions.
- Evaluation and validation of the proposed algorithms using simulation scenarios.

Requirements

- Knowledge of problems complexity, optimization algorithms and graph theory.
- Analytical thinking.
- C++

Contact

Dr. Aasem Ahmad <aasem.ahmad@tu-darmstadt.de>

References

[1] Kopanski, J., Rzadca, K.. *Plan-Based Job Scheduling for Supercomputers with Shared Burst Buffers*. In European Conference on Parallel Processin (pp. 120–135). Springer, 2021.

Issue date: **November, 2022**