Efficient Job Scheduling for Clusters with Shared Tiered Storage

Leah E. Lackner[1], Hamid Mohammadi Fard[*1], and Felix Wolf[1]

[1]Technische Universität Darmstadt – Germany

Abstract

New fast storage technologies such as non-volatile memory are becoming ubiquitous in HPC systems with one or two orders of magnitude higher I/O bandwidth than traditional back-end storage systems. They can be used to heavily speed-up I/O operations, an essential prerequisite for data-intensive exascale computing capabilities. However, since the overall capacity of the fast storage available in a system is limited, an individual job may not always benefit if access to fast storage implies longer waiting time in the queue. This is obvious if fast storage is shared across the system. We therefore argue that the decision of whether or not to use fast storage should be supported by the batch scheduler, which can estimate when the amount of fast storage a job desires will become available. We present a scheduling algorithm with this functionality and show in simulations significantly reduced makespan and turnaround times in comparison to always using fast storage, always using slow back-end storage, and random storage assignment.

Keywords: Scheduling, I/O, Resource management, Tiered storage, Non volatile memory, Data intensive applications

[*]Speaker