

Performance Estimation and Mapping of Structured Data to Parallel File Systems like Lustre

Contact:
Jakob Lüttgau
luetgau@dkrz.de

Abstract

Supercomputers use parallel distributed file systems to serve thousand of compute nodes with access to data. Data is spread across many data targets, which allows to read and write large files at high bandwidths. As part of this project/thesis you would develop a performance model and mappings to accounts for the configuration of the parallel file system. You will learn about low-level C programming to implement the mapping as well as advanced optimization options when using lustre. Lustre today serves seven out of the Top10 supercomputer in the world.