Parallel Computing and I/O

Distributed storage management policies in JULEA

- Author: Julian Benda
- Type: Bachelor's Thesis
- Date: 2022-03-27
- Reviewers: Jun.-Prof. Dr. Michael Kuhn, Kira Duwe
- Supervisors: Jun.-Prof. Dr. Michael Kuhn, Kira Duwe
- Download: PDF

A growing variety of storage media and technologies focus on hierarchical storage management and lead to more complex stora[ge] layouts. This work aims to develop a mechanism to implement HSM policies in JULEA to provide a streamlined framework that supports this type of research. JULEA is a storage framework for distributed systems, e.g. high-performance computing (HPC) systems. Implementing a modular mechanism to write high-level HSM policies enables extended object management on storage hierarchies, creating an effective environment for research. For constructing the HSM policy interface, different promising ideas f[or] HSM are analyzed. Based on this interface, the integration in JULEA is designed and implemented. To evaluate the performance various use cases and benchmarks were run with different simple policies, like LRU, and compared against JULEA without hierachy storage management (HSM). The qualitative analysis shows that the HSM mechanism reduces the overall performance for single tier scenarios up to 20%, due to additional metadata handling and function calls. However, comparing different policies shows that the performance of the lower storage tier can be completely hidden. In conclusion, the implemented mechanism work[s] as intended but should be revisited to minimize the performance impact. In addition, when using HSM, the application characteristics should be analyzed and used as the basis for selecting and parameterizing the policy.