
Performance Assessment of Metadata Management with Different Databases in a FUSE File System

- ▶ Author: Martin Barthel
- ▶ Type: Bachelor's Thesis
- ▶ Date: 2023-04-03
- ▶ Reviewers: Jun.-Prof. Dr. Michael Kuhn, Michael Blesel
- ▶ Supervisors: Jun.-Prof. Dr. Michael Kuhn, Michael Blesel
- ▶ Download: PDF

File systems not only hold the information inside a file but also information describing the entry, known as metadata. As several use cases like HPC raise the desire to make use of existing technologies able to be accessed from multiple machines and capable of managing large amounts of data, databases are a likely candidate as metadata storage solution. However, many categories of databases have been implemented over the years. Examples are key value stores, relational databases, graph databases and triple stores. This prompts the question which class should be picked. To provide arguments for that decision, an existing file system implementation is picked. This is provided by julea-fuse which already uses key value stores. julea-fuse is located within the JULEA-Framework. As the other metadata backend provided by it are relational databases, this is the obvious choice for the second database type to be evaluated. julea-fuse binds to the fuse library to support existing POSIX compliant applications while running in user space. To be able to support both backends via highly similar code an interface is created to hide their specifics. It is then evaluated whether julea-fuse completes common file system requests faster when using the key value store or the relational database. This is done via mdbench. There the key value store appears to be faster than the relational database across the evaluated methods. Hence the implied takeaway is that with the used file system structure, key value stores are the better storage backends for file system metadata in terms of latency.
