Metadata Services for I/O Middleware using MongoDB

Contact:
Jakob Lüttgau
luettgau@dkrz.de

Abstract

Large scale storage systems nowadays separate data and metadata because the access patterns are very different. Metadata is usually made up of many small I/O requests that result as a result of lookups or file system queries. Scientific applications use a wide range of domain specific metadata, that are used to catalogue and find simulation results or measurements. In this project/thesis, you would design and develop a metadata adapter for MongoDB to be used by I/O middleware. You would learn about data organisation and formats common in scientific applications (NWP/CLimate context) aswell as schemaless databases and replication.