L3S Research Center is investigating the interaction of people, applications, data, and systems on the web. The subjects of research are in particular very large data collections which originate from, e.g., crawling the web. These datasets are characterized by a high velocity, variety, and volume. Analyzing datasets that have sizes of several Terabytes requires special software, typically based on the MapReduce paradigm implemented by the Apache Hadoop framework, and special hardware that is particularly optimized for the storage and distribution of large datasets. The computing cluster of L3S is dedicated to such Hadoop-based computation and is the subject of this project.

Project abstract:
Advancing the L3S Hadoop Cluster

Project duration:
01.02.2014 - 31.10.2014

Members:
jaeschke
tunguyen

Project manager:
Prof. Dr. rer. nat. Robert Jäschke

Project research areas:
Intelligent Access to Information

Research Area:
Intelligent Access to Information

Status of the Project: