HDFS

Hadoop Distributed File System

Hadoop Distributed File System (HDFS) is designed to reliably store very large files across machines in a large cluster. It is inspired by the GoogleFileSyste.

General Information

- DFS_requirements. Summarizes the requirements Hadoop DFS should be targeted for, and outlines further development steps towards achieving this requirements.
- Documentation
  - HDFS Java API
  - HDFS Architecture Guide - a brief description of the design and architecture.
  - Hadoop_Upgrade. Describes a step-by-step procedure for manual transition of Hadoop cluster to a newer software version, and outlines enhancements intended to make the upgrade simple and safe.
- FAQ (look for the questions starting with HDFS.)
- HDFS_Publications