

The Large Hadron Collider (LHC) located at CERN, Geneva has finally been put in production, generating unprecedented amount of data.

These data are distributed across many computing centers all over the world that form the Worldwide LHC Computing Grid (WLCG).

One of the main issues since the beginning of the WLCG project is an effective file access on the site level in order to fully

exploit huge computing farms. The aim of this thesis is to explore existing data distribution work flows, standards, methods and protocols.

An integral part of the work is the analysis of jobs of physicists to understand input/output workloads and to discover

possible inefficiencies. Then, new upcoming solutions are evaluated in terms of performance, sustainability and integration

into existing frameworks. It is expected that these solutions will be based on distributed file systems such as NFS 4.1, Lustre and HDFS.