Distributed Data Management for the World's largest Machine

Sigve Haug

University of Oslo, Department of Physics, P. O. Box 1048, Blindern, 0316 Oslo, Norway
sigve.haug@fys.uio.no

Abstract. The worlds largest machine, the Large Hadron Collider (LHC) will have four detectors whose output is expected to answer fundamental questions about the Universe. The ATLAS detector is expected to produce 3.2 PB of data per year, which will be distributed to storage elements all over the world. In 2008 the resource need is estimated to be 16.9 PB of tape, 25.4 PB of disk and 50 MSI2k of CPU.
Grids are used to simulate, access and process the data. Sites in several European and non-European countries are connected with the Advanced Resource Connector (ARC) middleware of NorduGrid. This contribution reports on how ARCs data management, Globus Replica Location Service and ATLAS software are combined and used as a comprehensive distributed data management system. Observed limitations in the system and possible improvements are discussed.