

# Large File System Optimisation and User Education

*Tuesday 28 May 2013 17:40 (5 minutes)*

The storage volume in the CERN Computing Center is growing constantly and exceeded 100 PB in February 2013. To increase the efficiency of analysis tasks, the EOS storage system has been developed, which is optimized to handle random access to physics data. The current setup is running with a disk volume of 40 PB and more than 1000 users.

We analyse file related metrics such as throughput, read-ratio or reopen-ratio which can be obtained server-side. Based on first measurements of user-system interaction, we see multifaceted access characteristics. Inefficient file access can be identified such as up to thousands of re-opens and re-reads of a file. Inspired by database query improvement research, we have analysed the typical user access patterns to improve the system. This includes the modification of the software and configuration parameters such as the buffer size. In addition, it includes the improvement of the user access pattern as well. For that we want to provide additional usage and performance metrics which will help each user to validate the efficiency of their own code.

**Author:** ZIGANN, Philipp (CERN)

**Co-authors:** DUELLMANN, Dirk (CERN); Prof. BALKE, Wolf-Tilo (TU Braunschweig)

**Presenter:** ZIGANN, Philipp (CERN)

**Session Classification:** Lightning Talks