



Contribution ID: 273

Type: **Poster presentation**

Popularity Prediction Tool for ATLAS Distributed Data Management

Monday, 14 October 2013 15:00 (45 minutes)

This paper describes a popularity prediction tool for data-intensive data management systems, such as the ATLAS distributed data management (DDM) system. The tool is fed by the DDM popularity system, which produces historical reports about ATLAS data usage and provides information about the files, datasets, users and sites where data was accessed. The tool described in this contribution uses this historic information to make a prediction about the future popularity of data. It finds trends in the usage of data using a set of neural networks and a set of input parameters and predicts the number of accesses in the near term future. This information can then be used in a second step to improve the distribution of replicas at sites, taking into account the cost of creating new replicas (bandwidth and load on the storage system) compared to the gain of having new ones (faster access of data for analysis). The tool ensures that the total amount of space available on the grid is not exceeded. This information can then help to make a decision about adding and also removing data from the grid to make a better use of the available resources. The design and architecture of the popularity prediction tool is described, examples of its use are shown and an evaluation of its performance is presented.

Primary author: BEERMANN, Thomas (Bergische Universitaet Wuppertal (DE))

Co-authors: MOLFETAS, Angelos (University of Sydney (AU)); NAIRZ, Armin (CERN); SERFON, Cedric (CERN); SCHIKUTA, Erich (University of Vienna); STEWART, Graeme Andrew (CERN); Dr GOOSSENS, Luc (CERN); LASSNIG, Mario (CERN); BARISITS, Martin (CERN); VIGNE, Ralph (University of Vienna (AT)); GARONNE, Vincent (CERN)

Presenter: STEWART, Graeme Andrew (CERN)

Session Classification: Poster presentations

Track Classification: Data Stores, Data Bases, and Storage Systems