



Contribution ID: 487 Contribution code: THU 19

Type: Poster

Label-based virtual directories in the dCache storage system

Thursday 24 October 2024 16:00 (15 minutes)

Traditional filesystems organize data in directories. The directories are typically a collection of files whose grouping is based on one criteria, i.e., the starting date of the experiment, experiment name, beamline ID, measurement device, or instrument. However, each file in a directory can belong to different logical groups, such as a special event type, experiment condition, or a part of a selected dataset.

The dCache is a storage system developed to store large amounts of scientific data, used by many HEP and Photon Science experiments.

With recent developments in dCache, we have introduced a concept of file tagging, which dynamically groups files with the same label into virtual directories. The file labels can be added, removed, renamed, and deleted through the admin interface or via Rest API. The files in the virtual directories are exposed through all protocols supported by dCache.

This contribution will describe the details of the implementation for the file tagging in dCache and present our future development plans on automatic metadata extractions, a feature that will significantly simplify data management. Additionally, we are exploring the use of virtual directories as a way to translate scientific data catalogs into filesystem views for direct data analysis

Primary authors: Dr GREEN, Christopher (Fermi National Accelerator Lab. (US)); LITVINTSEV, Dmitry (Fermi National Accelerator Lab. (US)); CHITRAPU, Krishnaveni (National Supercomputer Centre in Sweden); MORSCHEL, Lea; SAHAKYAN, Marina; MEYER, Svenja; Mr MKRTCHYAN, Tigran (DESY)

Presenter: SAHAKYAN, Marina

Session Classification: Poster session

Track Classification: Track 1 - Data and Metadata Organization, Management and Access