



Contribution ID: 65

Type: **not specified**

Label-based Virtual Directories In dCache

Monday 31 March 2025 14:50 (20 minutes)

Traditional filesystems organize data into directories based on a single criterion, such as the starting date of the experiment, experiment name, beamline ID, measurement device, or instrument. However, each file within a directory can belong to multiple logical groups, such as a special event type, experiment condition, or part of a selected dataset. dCache, a storage system designed to handle large volumes of scientific data, is widely used in High Energy Physics (HEP) and Photon Science experiments. Recent advancements in dCache have introduced the concept of file tagging, which dynamically groups files with the same label into virtual directories. These file labels can be added, removed, renamed, and deleted through an admin interface or via a REST API. The files in these virtual directories are accessible through all protocols supported by dCache.

This presentation will delve into the implementation details of file tagging in dCache and outline our future development plans, including automatic metadata extraction. This feature aims to significantly simplify data management. Furthermore, we are exploring the use of virtual directories to translate scientific data catalogs into filesystem views, enabling direct data analysis. We will also discuss our new developments in the context of the National Analysis Facility (NAF) at DESY.

Desired slot length

Speaker release

Yes

Author: SAHAKYAN, Marina

Co-authors: Dr GREEN, Christopher; LITVINTSEV, Dmitry (Fermi National Accelerator Lab. (US)); HOYOS, Karen (DESY); MORSCHEL, Lea; Mr MKRTCHYAN, Tigran (DESY)

Presenter: SAHAKYAN, Marina

Session Classification: Storage & data management

Track Classification: Storage & data management