20th International Conference on Computing in High Energy and Nuclear Physics (CHEP2013)

Contribution ID: **425** Type: **Oral presentation to parallel session**

A modern web based data catalog for data access and analysis

Monday 14 October 2013 17:47 (22 minutes)

The SLAC Computing Applications group (SCA) has developed a general purpose data catalog framework, initially for use by the Fermi Gamma-Ray Space Telescope, and now in use by several other experiments. The main features of the data catalog system are:

- Ability to organize datasets in a virtual hierarchy without regard to physical location or access protocol
- Ability to catalog datasets stored at multiple locations and with multiple versions
- Ability to attach arbitrary meta-data to datasets and folders
- Web based and command line interfaces for registering, viewing and searching datasets
- A data "crawler" to verify catalog integrity and automate meta-data extraction
- A download manager for reliable download of collections of files

In this paper we will describe a recent project to update the data catalog to current web standards, in particular to:

- Isolate the database back-end from the server-side middle-ware by use of a file abstraction layer
- Develop Restful interfaces to make the server side functionality accessible to many tools and languages
- Develop a modern HTML5 based web client which also communicates with the server using Restful interfaces, and provides dynamic functionality such as drag and drop file upload/download.

These improvement open the way to integrating components of the data catalog with different back-end systems, and to provide a portal to support not only access to data, but to be able to operate on and analyze data remotely.

Authors: VAN KLAVEREN, Brian (SLAC); JOHNSON, Tony (SLAC)

Presenter: VAN KLAVEREN, Brian (SLAC)

Session Classification: Distributed Processing and Data Handling B: Experiment Data Processing, Data Handling and Computing Models

Track Classification: Distributed Processing and Data Handling B: Experiment Data Processing, Data Handling and Computing Models