



Contribution ID: 462

Type: **oral presentation**

Archiving Scientific Data outside of the traditional High Energy Physics Domain, using the National Archive Facility at Fermilab

Tuesday 14 April 2015 17:00 (15 minutes)

Many experiments in the HEP and Astrophysics communities generate large extremely valuable datasets, which need to be efficiently cataloged and recorded to archival storage. These datasets, both new and legacy, are often structured in a manner that is not conducive to storage and cataloging with modern data handling systems and large file archive facilities. In this paper we discuss in detail how we have created a robust toolset and simple portal into the Fermilab Archive Facility, which allows for scientific data to be quickly imported, organized and retrieved from the 0.650 Exabyte facility.

In particular we discuss how the data from the Sudbury Neutrino Observatory (SNO) for the COUPP dark matter detector was aggregated, cataloged, archived and re-organized to permit it to be retrieved and analyzed using modern distributed computing resources both at Fermilab and on the Open Science Grid. We pay particular attention to the methods that were employed to “uniquify” the namespaces for the data, derive metadata for the over 460,000 image series taken by the COUP experiment and what was required to map that information into coherent datasets that could be stored and retrieved using the large scale archives systems.

We describe the data transfer and cataloging engines that are used for data importation and how these engines have been setup to import data from the data acquisition systems of ongoing experiments at non-Fermilab remote sites including the Laboratori Nazionali del Gran Sasso and the Ash River Laboratory in Orr, Minnesota. We also describe how large University computing sites around the world are using the system to store and retrieve large volumes of simulation and experiment data for physics analysis.

Author: Dr NORMAN, Andrew (Fermilab)

Co-authors: Dr LYON, Adam (Fermilab); MENGEL, Marc (Fermilab); Dr DIESBURG, Michael (F); GHEITH, Michael (Fermilab); Dr ILLINGWORTH, Robert (Fermilab)

Presenter: Dr NORMAN, Andrew (Fermilab)

Session Classification: Track 3 Session

Track Classification: Track3: Data store and access