

## CHEP 2012 Abstract

Title: Taking Global Scale Data Handling to the Fermilab Intensity Frontier

Authors:

Robert Illingworth, Marc Mengel, Andrew Norman, Rick Snider, Adam Lyon (primary)

Fermilab Intensity Frontier experiments like Minerva, NOvA, g-2 and Mu2e currently operate without an organized data handling system, relying instead on completely manual management of files on large central disk arrays at Fermilab. This model severely limits the computing resources that the experiments can leverage to those tied to the Fermilab site, prevents the use of coherent staging and caching of files from tape and other mass storage media, and produces an onerous burden on the individuals responsible for data processing.

The SAM data handling system[1], used by the Fermilab Tevatron experiments CDF and D0 for Run II (2002-2011), solves these problems by providing data set abstraction, automated file cataloging and management, global delivery, and processing tracking. It has been a great success at CDF and D0 achieving global delivery rates of ~1.5 PB/week/experiment for raw data, Monte Carlo, production and analysis activities. However, SAM has been heavily tailored for integration in both CDF and D0 analysis frameworks, making it difficult and time-consuming to repeat that work for new experiments. The command line user interface is also complex, non-intuitive and represents a tall barrier for new and casual users. These issues have slowed the adoption of SAM by Intensity Frontier experiments. The Fermilab Computing Sector is improving SAM with a generic “deployment-less” HTTP based client for analysis framework integration and an intuitive FUSE[2] based user interface to permit universal adoption of SAM across the Intensity Frontier.

We will describe these solutions in detail, their technical implementation, and their impact on the adoptability of SAM for new experiments.

[1] <http://projects.fnal.gov/samgrid/>

[2] <http://fuse.sourceforge.net/>