

Monitoring use case: Data Management on the Data Storage System

- **Description:**

Fermilab is the Tier 1 Regional Center for CMS experiment. It will store peta-bytes of data using disk and tape storage systems. The performance of the storage system varies. The site administrators need to monitor the availability of disk and tape resources and the data usage information. They may re-distribute the data on different storage systems based on the data access frequency, data type and user priority.
- **Contact:**

Yujun Wu (yujun@fnal.gov)
- **Performance events/sensors required:**
 1. Storage system information: CPU, memory, available free storage, network utilization, OS configuration, how many clients are connecting, failure rate, etc.
 2. Data information: data type, size, data location, data clustering information, data mirror information, data permission (readable, writable, etc.)
 3. Data access information: read/write, frequency, duration, size, time, user info, lock info;
- **How the performance information will be used:**
 1. Identify the bottleneck of the storage system: may put the “hot” data into several copies or cache some data
 2. Re-distribute the data depending on the frequency of the access. For example, moving the frequently accessed data from poor performance system or tape system to high performance disk storage system.
- **Access needed:**

summary statistics, streaming of data and historical information
- **Size of data to be gathered**

Individual statistics will be small; archived log file will be large after years of monitoring
- **Frequency data will be updated**

Depending on the different storage systems and type of data. The log information for tape may have an upper limit of 5 minutes; for disk, every 1-2 minutes.
- **Frequency data will be accessed**

Will be accessed frequently.
- **Overhead constraints**

Daemons with a low overhead may need to be running periodically on the file servers.
- **Duration of the logging:**

TBD. Depending on the purpose of the data to be used, it may be kept several weeks or some summary information may be kept for several years.

- **Platforms:**

Requirement: Linux and Solaris, Disk Storage System, Tape Storage System (Enstore)