

BNL Batch and DataCarousel systems at BNL: A tool and UI for efficient access to data on tape with fareshare policies capabilities

Tuesday, 23 February 2010 15:15 (25 minutes)

The BNL facility, supporting the RHIC experiments as its Tier0 center and thereafter the Atlas/LHC as a Tier1 center had to address early the issue of efficient access to data stored to Mass Storage. Random use destroys access performance to tape by causing too frequent, high latency and time consuming tape mount and dismount. Coupled with a high job throughput from multiple RHIC experiments, in the early 2000, the experimental and facility teams were forced to consider ingenious approaches. A tape access “batch” system integrated to the production system was first developed, based on the initial OakRidge National Lab (ORNL) Batch code. In parallel, a highly customizable layer and UI known as the DataCarousel was developed in-house to provide multi-user fairshare with group and user level policies controlling the sharing of resources. The simple UI, based on a perl module, allowed to create user helper script to restore datasets on disks as well as had all the features necessary to interface with higher level storage aggregation solutions. Hence, beyond the simple access at data production level, the system was also successfully used in support of numerous data access tools such as interfacing with the Scalla/Xrootd MSS plugin back end, similarly the dCache back end access to MSS. Today, all RHIC and Atlas experiments use a combination of the Batch system and the Datacarousel following a 10 years search for efficient use of resources.

In 2005, BNL’s HPSS team decided to enhance the new features such as improve the HPSS resource management, enhance the visibility of realtime staging activities, statistics of historical data for performance analysis. BNL Batch provides dynamic HPSS resource management and scheduled read job efficiently while the staging performance can still be further optimized in user level using the DataCarousel to maximize the tape staging performance (sorting by tape while preserving fareshare policies).

In this presentation, we will present an overview of our system and development and share the findings of our efforts.

Summary

Efficient access to data on tape tool & UI interfacing with storage aggregator for coordinated, fair and policy driven data access.

Primary authors: Mr YU, David (BROOKHAVEN NATIONAL LABORATORY); Dr LAURET, Jerome (BROOKHAVEN NATIONAL LABORATORY)

Presenter: Mr YU, David (BROOKHAVEN NATIONAL LABORATORY)

Session Classification: Tuesday, 23 February - Computing Technology for Physics Research

Track Classification: Computing Technology for Physics Research