

Efficient Access to Massive Amounts of Tape-Resident Data

Thursday, 13 October 2016 12:15 (15 minutes)

Randomly restoring files from tapes degrades the read performance primarily due to frequent tape mounts. The high latency and time-consuming tape mount and dismount is a major issue when accessing massive amounts of data from tape storage. BNL's mass storage system currently holds more than 80 PB of data on tapes, managed by HPSS. To restore files from HPSS, we make use of a scheduler software, called ERADAT. This scheduler system was originally based on code from Oak Ridge National Lab, developed in the early 2000s. After some major modifications and enhancements, ERADAT now provides advanced HPSS resource management, priority queuing, resource sharing, web-browser visibility of real-time staging activities and advanced real-time statistics and graphs. ERADAT is also integrated with ACSLS and HPSS for near real-time mount statistics and resource control in HPSS. ERADAT is also the interface between HPSS and other applications such as the locally developed Data Carousel providing fair resource-sharing policies and related capabilities.

ERADAT has demonstrated great performance at BNL and other scientific organizations.

Secondary Keyword (Optional)

Storage systems

Primary Keyword (Mandatory)

Software development process and tools

Tertiary Keyword (Optional)

Primary author: YU, David (Brookhaven National Laboratory (US))

Co-author: Dr LAURET, Jerome (Brookhaven National Laboratory)

Presenter: YU, David (Brookhaven National Laboratory (US))

Session Classification: Track 6: Infrastructures

Track Classification: Track 6: Infrastructures