Efficient Access to Massive Amount of Tape-Resident Data
David Yu, Jerome Lauret

- BNL Mass Storage System serves 3 major experiments:
  US-ATLAS, STAR and PHENIX

- Randomly restoring files from massive amount of tapes degrades the read performance primarily due to frequent tape mounts, forwards and rewinds, resource availability under multi-project environment.

- Unorganized file submissions will fully optimized resource, causing resource either not being fully utilized or overwhelmed. Unmanaged drive resource may result to unpredictable wait time and interruptions

- And so ERADAT (Efficient Retrieval and Access to Data Archived on Tape) was born ...

1. Mount tape 1
   - Restore file A
   - Dismount tape 1

2. Mount tape 2
   - Restore file B
   - Dismount tape 2

1. Mount tape 1
   - Restore file C
   - Dismount tape 1

2. Mount tape 2
   - Restore file D
   - Dismount tape 2

1. Mount tape 1
   - Restore file E
   - Dismount tape 1

2. Mount tape 2
   - Restore file B

15 operations
9 operations
Status & Updates

Since, enhanced features, stability and scalability
  – ERADAT: priority staging, resource guarantees, resource sharing, drive-generation awareness, enhanced Web UI, multi-level real-time debug, advanced thread control, performance graphs & reports, control interface (OS agnostic), ...
  – DataCarousel: ability to control ERADAT restore policies, request auto-throttling, better error trap and re-submission, privileged user bypass, UI control of service

• Use-case with ERADAT: Randomly restoring 704 x 10 GB files out of 21 tapes, with 15 available drives
  – Using 15 job-queues, it took only 70 minutes. Average ~ 1.7 GB/s. Used 21 mounts.

• Data restored from tapes in 2015: ~17 PB, ~10 million files