BLUE WATERS

SUSTAINED PETASCALE COMPUTING

Nearline Data Retrieval

Brett Bode


GREAT LAKES CONSORTIUM
FOR PETASCALE COMPUTATION

Status

- The Blue Waters Nearline environment currently holds:
  - 35PB
  - 355 million files
- Nearline is currently scheduled to be turned off 12/19/2019 with the rest of Blue Waters!
  - Each science team is responsible for transferring your data to a different storage system!
  - NCSA does not know how much of the data remains to be transferred off site!
- There is a pending supplement to fund storage systems until 3/31/2020 to provide teams active until 12/19/19 the standard 90 day grace period for data transfer.

The good, the bad and the ugly

- For large files, 100GB or larger, data streams out off of tape at ~300MB/s or more.
- For small files – 16MB to 16GB – the rate is ~1-2 files per minute
- For tiny files <16MB the rate is ~10 per minute.

Tiny File Problem

- At ~10 files per minute the maximum you might expect to retrieve in the remaining life of nearline is ~300k files.
  - Multiple teams have millions of these files.
- NCSA is currently working on a way of utilizing space on the home lustre file system as cache space for HPSS to bulk load all small files into cache.
  - Once complete in mid-June this should make it possible to retrieve all of this class of file.
  - If you plan to put these files on tape somewhere else they will thank you if you tar them up first!

Small File Problem

- There is ~23PB in the small file class.
  - Much too large to solve by bulk loading to a disk cache.
- Tape count is also much larger. Some teams have data spread across more than 7000 tapes!
- Teams with lots of files in this category should already be transferring data!
- Talk to us, we are trying out a tool for prestaging files to disk cache that does improve that speed over Globus.
  - We need teams to be responsive for this to be effective!

Large Files

- Large files will stream off of tape via HPSS and Globus without any special effort by our team.
- Just make sure you leave plenty of time for the transfers and factor in lost time due to faults.

Data Repackaging

- Data repackaging is NOT required by NCSA for moving data off of nearline.
- However, the site you are moving data to may impose file count limits that force repacking data into larger tar files.
- If you no longer have computational access to Blue Waters we may consider a small allocation to allow data repacking. Please submit a ticket for the request.

Keys to Data Transfer Throughput

- Keep three transfers running all the time.
  - You can submit more than three, but only three will run at once.
  - If you want to prioritize certain transfers let us know, we can pause other transfers.
  - If you have other means of validating your data you might consider turning off Globus Checksums
- Monitor your transfers!!!
  - Credentials on both ends expire and need to be renewed!

Storing data on Nearline

- Nearline is still open for new data, but…
  - Requests for a quota increase will be closely reviewed and you must clearly state why the data should be on nearline (for up to six months) and where and when it will be transferred later!

- Ideally store data in 100GB or larger files.
  - No more than 1,000 files smaller than 100GB!

General requests

- Remove data after it is no longer needed on Nearline.
  - Do it now and following successful transfers!
  - Allows NCSA to have a better idea of the remaining problem.
  - Also the size of data, particularly in the range of 16MB to 16GB affects the viable performance improvement strategies.
  - You can perform deletes yourself or by opening a ticket.

General Requests

- Please respond to NCSA staff requests for information promptly!
  - Timely information is very important to efficiently managing the data retrieval problem.
- See https://bluewaters.ncsa.illinois.edu/data-retrieval-nearline for more information.
- If you have problems or questions please open a ticket! (email help+bw@ncsa.Illinois.edu)