A New 2PB GPFS file system for the JGI “projectb”

Jason Hick
Storage Systems Group
Lawrence Berkeley National Lab
10 February 2012
The new 2PB “projectb” file system is available on Phoebe now

- Some high level specs for users
- 2.8 PB
- Has a “per project” non-purged, backed-up portion with smaller quota (1-5 TB)
- And a “per user” purged, not backed-up portion with larger quota (10TB)
File systems best practices

• Unfortunately disk is still expensive

• All of the JGI’s data cannot continue to be stored on disk within the current budget

• Archive and delete data you no longer need

• Disk usage will be controlled through quotas in some cases and purging in others
There are two areas of storage within the “project” layout of the “projectb” file system.

- **Group directories**
  - Not purged
  - Subject to quota

- **User directories**
  - Purged
It is important for every group to come up with a data retention policy

How long should we keep the raw data?

Can the data be deleted or should it be archived?

Can we set up an automated way to archive and delete data?