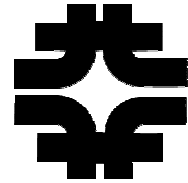

Astro: the SDSS Coadd

- Jim Annis EAG/CD
- An astrophysics application for OSG
  - relevant to
    - SDSS analysis
    - Dark Energy Survey R&D

Project Drivers & Scope

- The SDSS Coadd has
  - input
    - 3.5 TB, in the form of ~500,000 7 meg files
  - Jobs
    - ~ 1/2 hour, produce 10 TB of intermediate data
  - output
    - 0.5 TB, in the form of ~50,000 7 meg files

- The primary driver for us is the use of SRM
  - Secondary driver is non-FNAL, the metadata catalog of NOAO's NSA.

Project Milestones

- SDSS gsiftp server
  - Done. this was a great success, actually.
  - Campaign on access in Feb: > 40 threads a problem (we want >300)
- SRM to pre-place ~0.3 TB on usable grid sites
  - working this now
  - have done compute to dcache transfers
- direct transfer of intermediate files from grid worker nodes to dcache/enstore
  - soon
- SRM to position results back on EAG dcache server
  - soon, after we get the hardware

- We plan to start a campaign in May, 2005

Project Issues

- SRM to pre-place ~0.3 TB on usable grid sites
  - VDT SRM clients incompatible with CMS SRM servers
    - Have stopped work on current VDT SRM
    - Will use fnkits SRM
  - Where is srmReserve ?
    - ?!! No implementation of reservation in current SRM ???
      - see SRM Joint Functionality Design, v2.1.1 Final
        - » http://sdm.lbl.gov/srm-wg/doc/SRM.spec.v2.1.1.html
    - will have to human engineer this part
    - Or is the idea to have the Grid-admins edit the SRM config file?
      - this is nuts, of course
    - In May/June we are to use SRM sites, right? What do they do for reserving space?
      - I only need 3-6... (if 3, I need ~1 TB at each)
  - Could we have a Hands-On-SRM talk/class?

Effort Profile

- Neha Sharma, 0.5 FTE, working SRM and processing
- Mike Wilde and group at UC: R&D
- Jim Annis, Huan Lin, Chris Stoughton, scientists at work
  - (...hmm. We don't seem to be in the same league as the other projects)
- new hardware:
  - DES enstore account (done)
  - disk server as dcache front end (soon)
  - tam analysis clusters/OSG node (soon)

Risks

- Technical risks
  - SRM is feature poor at the moment, so use is driven by R&D
    - where is my srmReserve?
  - Grid generically doesn't like data intensive science

- Schedule risk
  - opportunity cost for science