GIP in OSG 1.0.0

Brian Bockelman

11 September 2008

GIP

- "Generic Information Provider"
  - The software used to advertise your site to the grid.
  - Used since OSG 0.1.6 (maybe Grid3?)
  - Started a major rewrite in OSG 1.0.0; will be finished in OSG 1.2.0

GLUE

- GIP uses the LDIF version of GLUE 1.3, a grid description language.
- GLUE allows you to describe site topology and configuration, as well as (some) realtime data.
- We are constantly working to get the GIP to be more true to the spec (it's hard!).

GLUE Layout

- Site: Top-level organization
  - Clusters: Entire cluster.
    - CEs: Roughly corresponds to a batch system queue, not an OSG install.
    - SEs: Entire storage element.
    - Services: Web services at the site.

GLUE Layout: Cluster

- Cluster: Top-level view of a cluster (headnode hostname, etc)
  - Subcluster(s): sets of non-overlapping heterogenous hardware descriptions and count.
    - Software (multiple): Installed software for each subcluster.

GLUE Layout: CE

- The GLUE CE is the "primary" object folks care about.
  - Should correspond to a queue (running jobs, queued jobs, policies, etc).
  - Has VO specific details in the VOView
  - Associated to a cluster.

GLUE Layout: SE

- SE: Totals about storage element.
  - SA (Storage Area): A non-overlapping logical grouping of storage. Corresponds to a link-group or poolgroup in dCache.
    - VOInfo: VO-specific view of a SA.
- Access endpoints: (gridftp/dcap/xrootd servers)
- Control endpoints: SRM server

GLUE Layout: Service

- A service is a generic web service running at a site.
  - New addition to GLUE 1.3; not that well thought through as you can't associate it to a cluster, CE, or SE.
- However, this is what we use to advertise SRM
  - Globus WS in the future?

GIP and GLUE

- It is the GIP's job to query all your local system resources (or at least read the config files) and write out the description of the site in GLUE.
  - Supports generic storage, BeStMan (OSG 1.2), and dCache.
  - Supports Condor, PBS, LSF (slightly broken in 1.0), SGE (slightly broken in 1.0).
    - SGE and LSF will work better in 1.2.

GLUE and the World

- Why is it important to understand and watch your GLUE output?
  - This is how the world looks at your site!
  - gLite WMS/RB for CMS sites
  - ReSS for OSG.
  - VOs use this to determine "how many batch slots are at site X?" and "what endpoints or batch system do I use at site Y?"

Why GLUE sucks

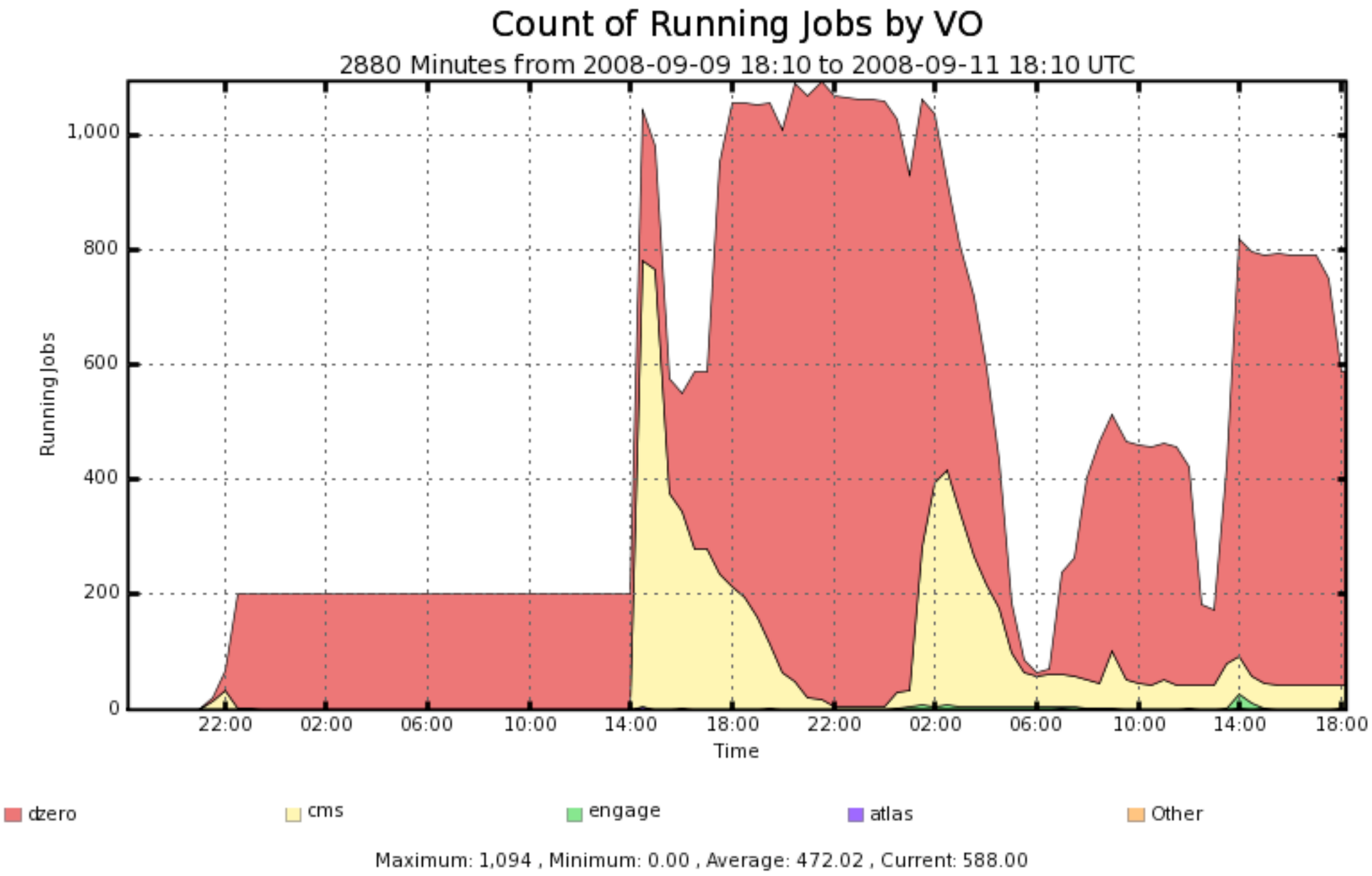
- GLUE can suck for the following reasons:
  - LDIF is a very, very brittle format and little validation is done; one bad attribute and the whole entry can be thrown out.
    - Hard to get things right: we have to program for every site's quirks.
- Code written to target OSG GIP, not GLUE 1.3. We were the IE5 of information providers.
  - Hence, we have the awesome job of bringing things back into compliance without killing

I watch your GLUE

- Every 5 minutes, I record the output of your CE's information and graph it.
- Next couple of slides are pictures of the status of Nebraska over 2 days.
  - Please use the URLs for your site!
  - Please let me know if they don't match your site!

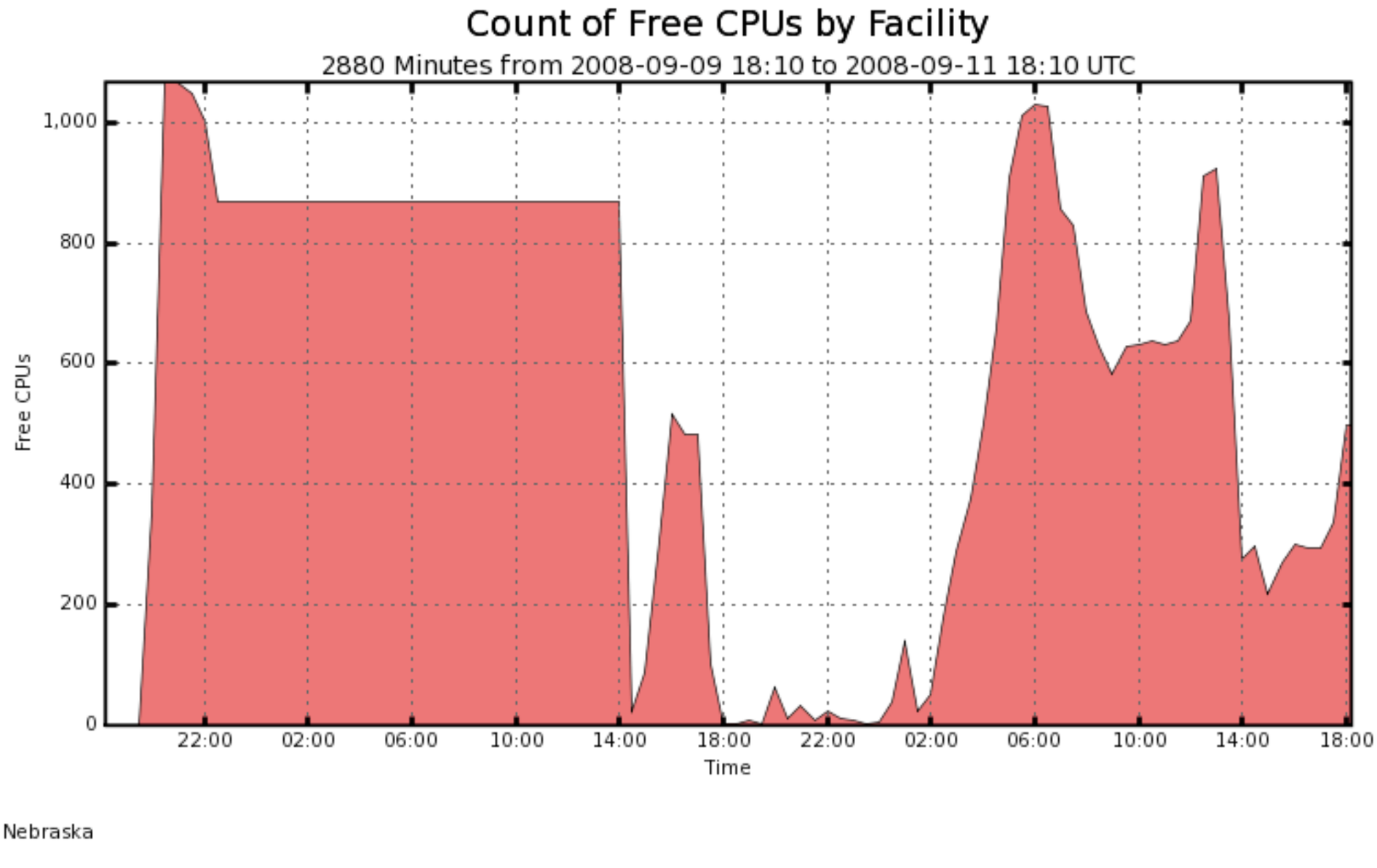
Nebraska Running

- http://t2.unl.edu/gratia/xml/gip_vo?facility=Nebraska



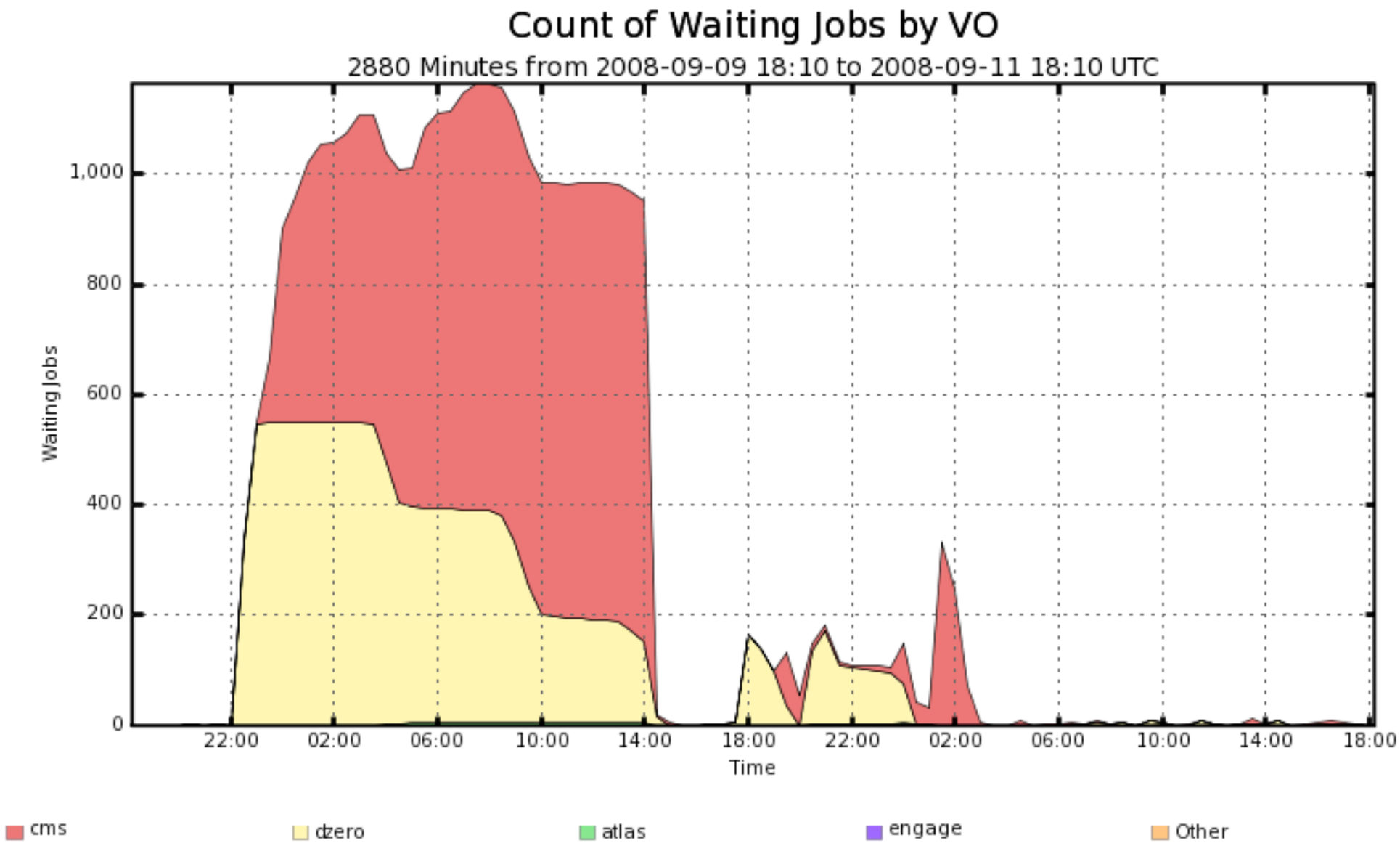
Nebraska Free Slots

- http://t2.unl.edu/gratia/xml/gip_free_cpus_history?facility=Nebraska



Nebraska Waiting

- http://t2.unl.edu/gratia/xml/gip_vo_waiting?facility=Nebraska



GLUE resources

- I'm a geek and I read the raw data:
  - http://is.grid.iu.edu/cgi-bin/status.cgi
- Developer Tests:
  - http://home.fnal.gov/~tiradani/xml_results/Other/index.xml
- Storage Viewer:
  - https://osg-ress-2.fnal.gov:8443/test/osg_storage.jsp