1.0.4 - LIGO Applications

Kent Blackburn
(Robert Engel, Britta Daudert)

Einstein@Home on the OSG

- No significant changes since last reported … see typical production running
- Continuing to see an average throughput of 50 to 100K CPU-Hrs per day. Discernable trend towards lower numbers over the past few weeks.
- Running on roughly 30 sites around the OSG
- This is a sweet-spot in terms of hands-on oversight and productivity.
- This is strictly opportunistic jobs
- Competition for cycles regularly impacts throughput, but nothing significant yet
- Detailed graphs at http://t2.unl.edu/gratia/vo?vo=ligo
- Einstein@OSG has been showcased numerous times in various OSG articles over the past few months

E@OSG Future Thoughts

- Some interest in developing support for SRM SE storage as opposed to the local storage currently used
- Probably greater interest in migrating to using a standard "pilot" method to replace LIGO/GEO home grown job submission/management solution
- Errors encountered have been documented in a report by Robert and given to OSG Production for improved usability for all communities.
- Unclear how much effort is available in the near term to make modifications/enhancements to E@OSG, but it is not a lot.

Binary Inspiral Workflows

- Focusing on several areas in an effort to achieve scientific contribution from the OSG
  - Use of storage elements to pre-stage (TBs) of data used by workflows designed by LIGO scientist
  - Evaluating data movement between LIGO data repositories and OSG SRMs
    - Details available at: http://www.ligo.caltech.edu/~bdaudert/INSPIRAL/FILE-TRANSFERS/STATS/
  - Changes to both LIGO data analysis codes and Pegasus to make the two grids appear transparent to the workflow submitters
  - Begun working with glide-in workflows (corral) to address some queuing competition originally seen between Binary Inspiral and E@OSG on Nebraska
- Almost all effort has taken place at two sites: Caltech/LIGO ITB cluster and Nebraska Firefly.
  - Looking at using other sites … hope to identify two new sites on production grid soon

Binary Inspiral Future Thoughts

- Currently studying performance of data movement and job runs on Firefly in comparison to runs on the LIGO Data Grid (no data movement on LDG)
- Special task-force looking into the usability of OSG opportunistic storage with LIGO.
- Exploring enhancements to Pegasus to support running on more sites and using glide-in technology.

Conclusions

- LIGO's E@OSG application is in science production contributing to the E@Home results. This is a huge analysis requiring enormous amounts of computation globally before conclusion at the level of a scientific publication are mature … its will happen!
- The need for large data sets available to the jobs in the binary inspiral workflows has been challenging based on the current usability of storage on the overall OSG "Grid".
  - Performance has yet to be demonstrated that matches well with the LDG for jobs and data throughput on the OSG
  - Long way from contributing scientifically in the binary inspiral analysis arena.