OSG Area Coordinators' Meeting

10.0.4 LIGO Applications (NEW)

Kent Blackburn

Caltech / LIGO

October 29th, 2009

Open Science Grid

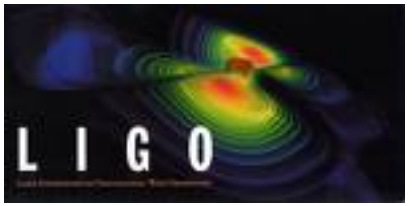
Current Initiatives

- LIGO Applications Coarsely fall into four groups:
  - Periodic Search Codes
    - (e.g., pulsars)
  - Compact Binary Coalescence Codes
    - (e.g., black holes, neutron stars, exotic stars)
  - Burst Search Codes
    - (e.g., supernovae)
  - Stochastic Search Codes
    - (e.g., primordial gravitational waves from early universe)
- Only the Periodic and Compact Binary Coalescence applications demand sufficiently large computing resources to warrant exploring non LIGO Data Grid resources!

Current Initiatives (cont)

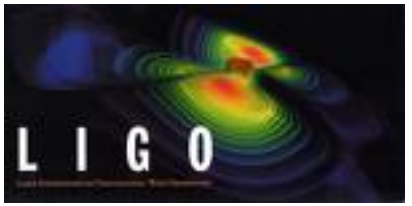
- LIGO Applications have existed for more than a decade before the LHC era
  - They are based on the LIGO Data Grid (LDG) architecture and its unique advantages to our scientific user community and analysis needs.
  - Migration onto the OSG is a challenge constrained by the requirement to preserve the integrity and trust of the LIGO Scientific Collaboration (60+ institutions, 650+ members)
  - Focusing on the low hanging fruit
    - Einstein@Home codes (BOINC Engine) for periodic applications
    - ihope / Pegasus workflow planner codes for compact binary coalescence applications
    - Both can gain scientifically from having additional resources beyond existing internal LIGO Data Grid resources

Key Accomplishments



- Einstein@Home on the OSG (E@OSG)
  - Close to 18 months of development have gone into this effort
  - Based on original code established for the German D-Grid
  - Began as a WS-Gram submission application
    - Since evolved to GT2 and now to Condor-G
- Currently E@OSG is contributing number two in the world on a weekly basis (excludes contribution made using non-gatekeeper submission mechanisms)
- Robert Engel will contribute approximately 25% of his time to maintaining and growing this contribution.
- Useful URL:
  - http://boincstats.com/stats/user_graph.php?pr=einstein&id=282952

Key Accomplishments



- Ihope compact binary coalescence on the OSG
  - This code base is maintained by the largest sub-community within LIGO
    - Highly dynamic code base
    - Highly optimized to fit within existing LIGO Data Grid resources (LDG resource scope is based on this application)
    - Constructed as a large workflow (~100,000 jobs)
    - Each workflow requires tens of terabytes of LIGO data be readily available
    - All effort to date has been prototype / porting activities on small none interesting segments of the data
    - Pegasus is adopted/accepted by this sub-community for workflow management of ihope
  - Britta Daudert will be working close to 100% on the evolution of this application to bring OSG into the arena of science contributors.
  - Not a production application for the OSG at this time. LIGO is successful with existing LDG resources.

Key Issues and Obstacles

- Einstein@Home on the OSG
  - Going very well as a production application
  - The amount of storage needed for this application scales with the number of worker nodes at a site that are participating.
    - A few sites have insufficient storage in the $OSG_DATA area to allow full utilization.
    - Miron has offered to pay for additional storage (order Terabyte) where this is a problem.

Key Issues and Obstacles (cont)

- Ihope application on the OSG
  - Lots of obstacles revolving around the large volumes of data needed to run a scientifically valued workflow!
  - OSG's own views of its Storage Solution are only now emerging
    - Finding a unified, transparent "view" of using large storage resources on the OSG is a challenge
    - Need a solution that is "automatic". Currently much work has to be done by hand to move from site to site on the OSG.
  - Evolution requires change to LIGO and Pegasus base codes
    - These changes burn up good will, trust and manpower to keep up with evolution (e.g., Gridcat -> VORS -> MyOSG -> OSG-MM)
  - Opportunity emerging to gain additional support from Syracuse University under the LIGO domain
    - Will benefit from having a well defined storage architecture in place in the OSG for opportunistic compute/storage applications.

WBS Exception Reporting

- Currently finishing up a year three milestone to prove that a large LIGO workflow (with 50,000 jobs) can use SRM / OSG SE technology.
- This workflow has been running for close to a month on the Caltech ITB cluster which has over 100 worker nodes.
- Due to an issue with condor (which has been reported) only a few jobs are actually being run on this large resource and what was expected to complete in a week is taking a month.
- In principle this has been accomplished.