

VO Experiences with Open Science Grid Storage

OSG Storage Forum | Wednesday September 22, 2010 (10:30am)

LIGO

BRITTA DAUDERT & KENT BLACKBURN

LIGO LABORATORY | CALIFORNIA INSTITUTE OF TECHNOLOGY

BACKGROUND ON LIGO'S BINARY INSPIRAL WORKFLOW

- ✘ LIGO instruments generate more than a TB of data each day
- ✘ Roughly 15% to 20% is analyzed in near real time
 - + This data is “replicated” in near real time to all LIGO Data Grid sites
 - + Data management is in the domain of the LIGO Data Grid, not the application!
- ✘ Binary Inspiral Workflows contain many 10s of thousand of jobs just to analyze a few weeks of data
 - + Workflows (and architecture of LIGO Data Grid) designed around analysis of weeks of data at a time
- ✘ Remaining data relevant to time intervals of scientific significance
- ✘ LIGO will finish the second science run this October and will begin an extensive upgrade to the Advance LIGO instrumentation
 - + ~10x more sensitive
 - + ~ 1000x larger window on the universe and corresponding event rates
 - + Workflows design may be significantly larger in the future or it may be drastically redesigned to support LIGO data in an open data era

BEFORE OSG STORAGE ELEMENTS

- ✘ Data were transferred during execution of the binary inspiral workflow
 - + Pegasus optimized this for storage minimizations, not computation
- ✘ One dag node was dedicated to transfer the entire data set of a work-flow into \$OSG_DATA
- ✘ The work-flow is executed in \$OSG_DATA since most dag jobs need direct access to the data
 - + Space in \$OSG_DATA very limited
- ✘ No OSG site prior to useful SE appearing on OSG were able to cope with scientific scale workflows of the binary inspiral
- ✘ Early SE to appear on the OSG did not allow LIGO VO to run there OR required changes to the scientific codes and execution to run ... unpleasant to think about for the science review committees!!

UTILIZATION OF OSG STORAGE ELEMENTS

- ✘ LIGO data set is pre-staged into an OSG storage element
 - + Britta takes care of this with custom scripts that make use of Britta being a member of the LIGO Scientific Collaboration ... access restrictions
- ✘ During execution of the work-flow, the data is sym-linked to the execution directory
 - + Pegasus workflow planner takes care of doing this
 - ✘ If the mapping between file spaces is known
 - ✘ SE file system needs to be mounted on all worker nodes
 - ✘ In the future we hope Pegasus functionality will expand to support “second level staging” from the SE to the worker nodes ... only data needed by jobs running on the work node are “staged” from the SE to the worker node

CHALLENGES AND EXPERIENCE ...

- ✘ Transferring LIGO data:
 - + What data sets need to be transferred is highly dependent on the specifics of a particular workflow ... as storage has to be “leased” with fixed size and life expectancy there is currently the requirement to be precise in specifying the data transfers based on the specific workflow
 - + File transfer and md5sum checks submitted from an OSG submit host. File transfers are executed locally on the OSG submit host via third party transfers. Md5sum checks are executed remotely on the OSG CE that is associated with the SE (competition for slots means this can take a very long time)
 - + RLS catalog at Caltech populated with information about files in the OSG SE ... this is needed by Pegasus workflow planner
 - + At present both globus-url-copy and srm-copy are being used, with performance issues in transfers noticed
- ✘ Britta has posted details of what is involved in the transfer at:
 - + <http://www.ligo.caltech.edu/~bdaudert/INSPIRAL/FILE-TRANSFERS/>

CHALLENGES AND EXPERIENCE CONTINUED

- ✘ To date, only a handful of OSG Storage Elements have been able to meet the LIGO requirements for this application:
 - + LIGO's ITB, Nebraska's Firefly, Caltech's CMS Tier 2 have demonstrated working prototypes of analysis
 - + UCSD's Tier 2, TTU_ANTAEUS and UMissHEP are in progress
- ✘ Not all of these are known to permit science scale workflows, but are providing useful experiences with using SE technology with this application
- ✘ Britta has posted statistics of testing at:
 - + <http://www.ligo.caltech.edu/~bdaudert/INSPIRAL/FILE-TRANSFERS/STATS/>

PERFORMANCE IMPROVEMENTS WITH GLIDE-IN

- ✘ We have successfully tested several workflows utilizing the Glidein tool Corral developed along side Pegasus
- ✘ The Glidein tool has also been tested on file transfer dags
- ✘ Use of Glideins may improve performance of LIGO Inspiral work-flow on the OSG since a number of Inspiral jobs are of very short duration (at least while prototyping)
- ✘ Glideins may also improve performance of transfer dags since md5sum jobs may sit in remote queues for long times at busy OSG sites
- ✘ Britta has metrics for LIGO experiences with glideins at:
 - + <http://www.ligo.caltech.edu/~bdaudert/INSPIRAL/GLIDEINS/STATS>

CONCLUSIONS

- ✘ LIGO's binary inspiral application needs large amounts of storage closely coupled to the worker nodes in its present form
- ✘ Few OSG sites with storage elements have been available to bootstrap this to a viable production scalable solution that meets LIGO's science publications environment
- ✘ A special task force has been pulled together to understand the challenges and develop the right tools and architecture to bring this to the next level.
- ✘ ...THE END