

Martin Sevior martines@unimelb.edu.au

## High level grid frameworks



# Very high level overview of:

- Bare DIRAC
- GANGA used by ATLAS,LHCB
- pathena used by ATLAS
- ddm, dg2 data distribution tools (ATLAS)
- CRAB CMS Remote Analysis Builder (Apologies to experts in the audience. Feel free to correct!)





Overview paper http://doc.cern.ch//archive/electronic/egee/tr/egee-tr-2006-001.pdf



- Glite provides the framework and Job Control Language (JDL)
- High Level middle-ware automates JDL production.
- Provides additional tools and services
- Provides monitoring





- •Tight integration with MetaData
- •Automatic failure recovery
- •Sophisticated Project progress information







**DIRAC** scripts

#### The execution: Phase I CPU days consumed by simulation Experiment / hour 12 Days from 2010-04-13 to 2010-04-25 80 70 250 VM running 60 20 VM running 50 days 40 **10 VM running** 20 10 0 2010-04-15 2010-04-17 2010-04-19 2010-04-21 2010-04-23 2010-04-25 Max: 79.49, Average: 7.79, Current: 0.32 e000049 71.6 e000045 20.9 Generated on 2010-05-01 15:23:48 UTC IBERGRID'2010, 24-27 May, Braga (Portugal) 13



**DIRAC** scripts

## Combined GRID, EC2 and local clusters





- Powerful and stable
- Scales very well
- Modular design
- Already backends for GRID, EC2 and local
- Job recovery
- Excellent web interface for job monitoring
- Fantastic developer support.
- Metadata integration?
- Progress Bar?





- Distributed data analysis tool
- Employed by ATLAS, LHCb and other experiments.
- Integrates with DIRAC, PAnda and "bare" grids (EGEE, OSG, NorduGrid)
- General solution with many options





GANGA

#### Run the default job locally: Job().submit() Default job on the EGEE grid: Job(backend=LCG()).submit() Listing of the existing jobs: jobs Get help (e.g. on a job): help(jobs) Display the nth job: jobs(n)

#### Copy and resubmit the nth job:

jobs(n).copy().submit()

#### Copy and submit to another grid:

- j=jobs(n).copy()
- j.backend=DIRAC()
- j.submit()

#### Kill and remove the n<sup>th</sup> job:

job(n).kill()
job(n).remove()



## CLI, scripting and GUI interfaces





### Panda - ATLAS



- PanDA (Production and Distributed Analysis) is a pilot job system
  - Executes jobs from the ATLAS production system and from users
  - Brokers jobs to sites based on available compute resource and data
    - Can move and stage data if necessary
- Triggers data movement
   back to Tier-1s for
   dataset aggregation



THE UNIVERSITY OF **PATHENA** 

- Grid interface to the ATLAS Athena analysis package (essentially talks to Panda servers).
- Test setup on a single file locally
- Run complete on the grid with exactly the same command, except for full data set.
- Automatic matching of cpu+data resources
- Automatic job recovery (rerunning will only execute failed jobs)
- pathena preferred by Australian ATLAS students



- ddm service to move data files to preferred sites.
- Data transfers subject to authorizations above trigger threshold
- dq2 service used to retrieve root files from remote grid sites



- CMS Remote Analysis Builder
- CMS tool to create JDL to integrate with a collection of higher level services
- High level services similar to those provided by Panda and Gaudi
- Provides Data discovery
- Packs local user lib and configs
- Communicates with CRAB server



**CRAB** services





- Panda, CRAB server and DIRAC all provide similar functionality.
- Very similar to our project server
- pathena and CRAB are very specific solutions for ATLAS and CMS
- GANGA is a general solution available for ATLAS and exclusively used by LHCb
- Unimelb students prefer pathena
- Tight integration with analysis framework is a big drawcard.



- The data migration tools are essential
- dq2 essential to retrieve root files
- ddm is available to users with authorization.



- Worthwhile to continue our experiment specific interface for distributed resources
- Similar functionality provided by LHC experiments
- Reuse of DIRAC components could significantly speed development of our solution
- More details of both coming up :-)



© Copyright The University of Melbourne 2008