The GridPP DIRAC project

DIRAC for non-LHC communities

GridPP
UK Computing for Particle Physics

Imperial College London

- GridPP is a collaboration of particle physicists and computer scientists originally proposed to support the LHC experiments but it now non LHC experiments and researchers from other disciplines.

- LHC experiments have moved or are moving away from EMI WMS. DIRAC is being considered as a replacement for EMI WMS

- DIRAC also offers a file catalog – possible replacement for the LFC still widely used by small VOs.

- We have set up and are maintaining a DIRAC instance at Imperial College London.

Imperial College London

The first installation was done in April 2013. We are currently supporting following VOs:

- na62.vo.gridpp.ac.uk
- t2k.org
- snoplus
- cernatschool.org
- comet.j-parc.jp
- pheno
- northgrid
- londongrid
- gridpp

The GridPP DIRAC project

Dirac offers many services, some of them are particularly attractive for small VOs:

- WMS with pilot jobs
- DFC – file and metadata catalogue. LFC is just a file catalogue (+ a comment field…)
- Synchronous and asynchronous file transfers integrated with the DFC
- Can submit jobs to CREAM and ARC CEs and to cloud and VAC sites

- NA62 use the Grid to run their MC jobs.
- Output files are to be moved to central Castor storage at CERN and RAL
- We have a custom written job submission and monitoring system which allows shifters to distribute jobs over the Grid and monitor their progress.
- The submission can be done either to EMI WMS or GridPP DIRAC service
- The output files are copied to CERN and RAL using a File Transfer Controller (a Web Service based FTS client). This decouples the MC simulation task from data movement. We are planning to replace this service by DIRAC asynchronous file transfer service (e.g. use dirac-dms-replicate-and-register-request). Need to upgrade DIRAC to allow this.
- Also plan to replace the LFC by the DFC.

http://cernatschool.web.cern.ch/

From the website:

"CERN@school is a programme for school students - and driven by school students - that brings technology from CERN into the classroom to inspire the next generation of scientists and engineers".

They used the GridPP DIRAC instance to implement an example workflow:

- Upload a dataset
    1. Raw data from the CERN@school detectors;
    2. Add metadata to the dataset using Python API

Imperial College London

- Process the dataset on the Grid
    1. Select data files of interest using metadata query
    2. Run CERN@school software (distributed via CVMFS) on selected data
    3. Write output to a selected Storage Element
    4. Add metadata to the generated data.
- Run an analysis on the processed data
    1. Select data of interest using a metadata query
    2. Retrieve output from the grid based on the selection

    Work coordinated by Tom Whyntie from QMUL.

The GridPP DIRAC project

Imperial College
London

- Conservative DIRAC version updates to maintain balance between new features and stability
- DIRAC registry updated manually:
  1. defining new users and groups
  2. adding new sites and activating/deactivating them
  3. adding site resources (CE, SEs)
- We are working on automating this at Imperial (see the GridPP DIRAC Project talk on Thursday)
- We use a mailing list to give support to our users.
- Wiki: https://www.gridpp.ac.uk/wiki/Dirac

We periodically run SAM tests to check if sites are accepting jobs:

GridPP DIRAC SAM Tests

This monitoring is still experimental, and doesn't take into account whether sites are properly registered in DIRAC, or are in downtime.

The GridPP DIRAC Portal is very useful for looking at sites and jobs in detail.

Site	Last job run
LCG.Bristol.uk	2015-04-09 12:21:00
LCG.Cambridge.uk	2015-04-09 11:11:02
LCG.Durham.uk	-
LCG.Glasgow.uk	2015-04-09 11:12:50
LCG.Lancashire.uk	2015-04-09 11:11:26
LCG.Liverpool.uk	2015-04-09 11:52:47
LCG.Manchester.uk	2015-04-09 11:56:05
LCG.Oxford.uk	2015-04-09 14:18:13
LCG.RAL-HEP.uk	2015-04-09 11:13:35
LCG.Sheffield.uk	2015-04-09 11:53:07
LCG.UKI-LT2-Brunel.uk	2015-04-09 11:11:42
LCG.UKI-LT2-IC-HEP.uk	2015-04-09 11:10:54
LCG.UKI-LT2-QMUL.uk	2015-04-09 07:27:07
LCG.UKI-LT2-RHUL.uk	2015-04-09 11:11:10

Site	Last job run
VAC.BHAM-HEP.uk	2015-04-09 05:48:31
CLOUD.CERN.ch	2015-04-05 23:27:46
DIRAC.RAL.uk	2015-04-09 11:12:33
VAC.Lancaster.uk	2015-02-16 17:11:45
VAC.Manchester.uk	2015-04-09 11:18:49
VAC.Oxford.uk	2015-04-09 12:42:04
CLOUD.UKI-LT2-IC-HEP.uk	2015-04-09 11:07:59

Thanks to Andrew McNab

University of Manchester