

# LHCb user experiences on the Grid

Greig A. Cowan

University of Edinburgh

GridPP21: 3rd Sept 2008



- User analysis of LHCb data takes place at T0/T1.
- Jobs that do not depend on data can also go to T2.
  - Toy studies using ROOT/RooFit.
  - Generator level studies.
- In addition, institutes have their own (T3) compute/storage resources.
- Using Ganga as your interface makes it simple to switch between running on the Grid and running on CERN LSF or other batch system.
  - See presentation by Mark Slater yesterday.



## Things were **bad** for LHCb users on the Grid

- Job success rate was low and users were becoming disillusioned.
  - **Data access** was commonly cited as being responsible.
  - Incorrect **site software installation** was another.
  - Ganga often not picking up all user libraries correctly, leading to job failures.

## At Edinburgh...

- Users stopped running jobs on the Grid.
- Data was replicated to our Grid SEs and jobs were submitted to Condor/SGE batch systems (using Ganga).
  - Success rate was much higher.
  - This was helped by having **someone on site** who could run the storage and who actually used the system, so understood users problems.



## Things were **bad** for LHCb users on the Grid

- Job success rate was low and users were becoming disillusioned.
  - **Data access** was commonly cited as being responsible.
  - Incorrect **site software installation** was another.
  - Ganga often not picking up all user libraries correctly, leading to job failures.

## At Edinburgh...

- Users stopped running jobs on the Grid.
- Data was replicated to our Grid SEs and jobs were submitted to Condor/SGE batch systems (using Ganga).
  - Success rate was much higher.
  - This was helped by having **someone on site** who could run the storage and who actually used the system, so understood users problems.



## The Good

- Instabilities in the system have reduced and success rates are higher.
  - One user submitted  $\sim 1500$  jobs in 4 days with  $\sim 100\%$  success.
  - Users find the return time of jobs much faster than using CERN LSF.
    - This would indicate that LSF is still being used by a lot of people.
- Ganga 5 has been released.
  - Solved the user lib discovery problem (now using CMT).
  - Support for python job options  $\Rightarrow$  typos found prior to job submission.
- Users seem happy, but should they be with an 80% success rate?

## The Bad

- Data access continues to be cited as a problem (T1's only).
  - Partly due to problems with the site SE (dcap access not working).
  - Partly due to problems with the LHCb bookkeeping.
- DIRAC2 still managing user analysis - no longer supported by devs.
- Software evolving quickly, but the documentation is not (common prob).

## The Good

- Instabilities in the system have reduced and success rates are higher.
  - One user submitted  $\sim 1500$  jobs in 4 days with  $\sim 100\%$  success.
  - Users find the return time of jobs much faster than using CERN LSF.
    - This would indicate that LSF is still being used by a lot of people.
- Ganga 5 has been released.
  - Solved the user lib discovery problem (now using CMT).
  - Support for python job options  $\Rightarrow$  typos found prior to job submission.
- Users seem happy, but should they be with an 80% success rate?

## The Bad

- Data access continues to be cited as a problem (T1's only).
  - Partly due to problems with the site SE (dcap access not working).
  - Partly due to problems with the LHCb bookkeeping.
- DIRAC2 still managing user analysis - no longer supported by devs.
- Software evolving quickly, but the documentation is not (common prob).

## Likely issues

- DIRAC3 workload and data management now used in production. . .
- . . . but not yet being used to manage user analysis.
- Expect teething troubles when it all starts.

## Potential solutions

- Grid operations team are an expert group with contacts to users, sites and developers.
- New job monitoring tools in DIRAC3 will make it a lot easier for users to see what it going on.
- Improvement in Ganga should mean that users spend less time debugging problems with their own jobs.
- DIRAC3 comes with new and improved bookkeeping.