Phenogrid Status

David Grellscheid

IPPP, Durham University

david.grellscheid@durham.ac.uk

16th GridPP Collaboration Meeting
QMUL, London
2006-06-28

Outline

1. Main objectives
2. Herwig++ Genser integration
   - Herwig++
   - Genser
3. Phenogrid VO
   - Positive developments
   - Negative experiences
   - User interest

Main objectives

Phenogrid

http://www.phenogrid.dur.ac.uk/

has two main objectives

- Development and integration of *Herwig++* event generator into LCG Generator Services (Genser) software package
- Support UK phenomenology & theory community in Grid use: exploration of needs, not production service in first instance
  no monolithic use case!

Herwig++ overview

Herwig++

is a Monte Carlo Event Generator based on the ThePEG framework.
Monte Carlo EGs

- aim to simulate the full physics of a particle collision,
- are the starting point of the LHC simulation chain,
- provide the input for the detector simulations.

- Following a January beta release with LHC physics support, we now aim for regular stand-alone releases about every 6 months.
- Build procedure is standard GNU `configure; make; make install`
- all physics happens in replaceable libraries which are dynamically loaded on demand by the running executable

Genser integration

Genser

- is part of LCG Applications area,
- provides one-stop-shop for event generators,
- but is not included in any LCG SW by default!

March 2006 saw first Genser release containing Herwig++.
Had problems in implementation:

- build process of upstream tarball not used and removed
- all code (infrastructure and physics) linked into one non-dynamic library
- Main problem: very little communication with Genser.
- Try again in July, meeting Genser maintainers directly in CERN

Positive developments

- List of CEs that have *pheno* enabled. Thanks!

  `ce01.tier2.hep.manchester.ac.uk`
  `ce1-gla.scotgrid.ac.uk`
  `ce1.pp.rhul.ac.uk`
  `ce.epcc.ed.ac.uk`
  `dgc-grid-40.brunel.ac.uk`
  `fal-pygrid-18.lancs.ac.uk`
  `gw-2.ccc.ucl.ac.uk`
  `gw39.hep.ph.ic.ac.uk`
  `heplnx201.pp.rl.ac.uk`
  `lcgce01.gridpp.rl.ac.uk`
  `mars-ce.mars.lesc.doc.ic.ac.uk`
  `t2ce02.physics.ox.ac.uk`
  `valor del bdii`

- `g77` is on VO ID card, but no automatic mechanism to get it included. Maybe look into SFT for *pheno*?
- Storage transition to LFC went smoothly
- next: need VOMS before Dec'06
  (end of LDAP-based authentication)

Thank you to RAL support!

Negative experiences

LCG command line utilities very frustrating to use. Two examples:

Handling of logical file names

- one file can be replicated on several SEs
- lcg_utils allow file to be specified via LFN
- available replicas are listed in a certain (fixed) order
- if first entry's CE has a problem / is decommissioned
  $\Rightarrow$lcg_utils stall (no timeout, no error message)

- first experience with GGUS
- immediate first response from our LFC (Nikhef), made initial contact with developers
- developer response entirely unsatisfactory

Response from lcg_utils

- ```
  Sorry, but lcg_utils cannot be considered as "a buggy tool",
  because the replica retries is not supported.
  This is a deliberate choice : lcg_utils is aiming to be a generic
  *simple* low level tool, that can be used by different VOs/
  applications, no matter what they do - complex or not.

  Your user could just write a script looping on the existing replicas.
  ```
- bug-free by definition, but LFN concept now useless
- *"let the VOs do all the work"* approach does not work for us, we rely on lcg_utils. VO users have wide range of interests, cannot provide isolating layer on top of LCG that will suit everyone.
- What can be done? As far as developers are concerned, issue is closed.
  ```

Response from lcg_utils

- `Shouldn't the non existing physical files be deleted from the catalog before anything else?`
  `A storage element should not be removed without the catalog being updated, in my opinion...`
- Need procedure for SE shutdown
- right now: I have to monitor EGEE broadcasts, check if files are affected, and remove the replica entries from LFC by hand

VO scripts break

lcg-infosites

Plain text output of active resources for a VO. Used in script for SW installation, functionality check.

- Command line flags counter-intuitive:
- `--foo` vs. `-foo`
- depend on placement order: `--vo pheno -v 1` vs. `-v 1 --vo pheno`
- Behaviour changes from version to version:
- default output very verbose
- summary output `-v 1` broke between 2_4_0 and 2_7_0, was fixed
- now have debug output `valor del bdii` on stdout

All these may seem minor, but they are the difference between using the Grid and going elsewhere!

User interest

So far have a handful of users. More have expressed interest, why aren't they using Phenogrid?

Non-coding users:

- need to use software that is not licenced on all sites (NAGlib, Mathematica). Can GridPP help here?
- have no local UI available. gLite3 solves this, UI installation instructions on Phenogrid website
- 1000 trivially parallel jobs take 90 minutes to submit (RB is bottleneck)
- don't think about possibilities large-scale computing can offer, happy to run calculations on their desktops over the weekend

User interest

Software developers:

- development code changes frequently, VO SW area not suitable:
- SW installation toolset difficult to automate, needs lot of intervention, does not scale well
- stick with clusters they have locally, with direct access to home directories

Near Future

- Genser discussions on Herwig++ integration
- in Durham all new PhD students will be signed up to Phenogrid $\Rightarrow$ get feedback from larger user base
- individual follow-up with theorists who have expressed interest initially
- advertise Grid among theorists, need to give examples of what can be done
- Consider Portal again