LHCb's first steps in volunteer computing

Federico Stagni, on behalf of the LHCb distributed computing team

LHCb

Credits

- Baptiste Cabarrou - stagiaire, summer 2013
  - did the majority of the work

(in **alphabetic** order)

- Philippe Charpentier
- Michele Ferrari
- Mario Ubeda Garcia
- … and some testers

LHCb

Why volunteer computing?

- It's free
- And also not so difficult to set up, in a first approximation

Getting started for your VO

CERN IT Department

- Identify suitable applications
  - High CPU, small data applications are a natural match → **We run standard LHCb MC jobs**
- Native applications running on Linux only, or Virtualisation?
  - For cloud applications, you can use BOINC/Volunteer computing simply to dispatch VM images and then use your pilot framework to control the VM → **We did the work for clouds, this was fitting**
- Target resources?
  - Desktop grids and computer centre farms → **Trying that**
  - Volunteers among the general public → **We are not there, not yet at least**
    - The latter requires effort in terms of outreach and interaction

How it works

Just copying

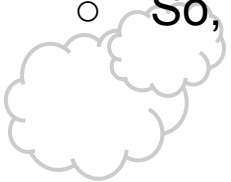
- Uses a virtual machine on the volunteer computers
  - User installs Virtual Box
  - User installs the BOINC client
    - Attach to project and ready to go with application on CernVM
    - The BOINC client downloads vboxwrapper that gets the image and the job to run on the VM.
    - Successful migration from CERN wrapper to BOINC native “Vboxwrapper' this year.

I stole (also) the above from Nils’s presentation: we did exactly the same as T4T did!

How it works/2

- Server up and running: http://lhcbathome.cern.ch/Beauty/
  - The BOINC application is the VM itself
  - Got good suggestions from Test4Theory group
  - Configuration isn't easy, need to experiment more
  - Out of the box integration within LHCbDIRAC
    - **To see the first jobs running, we didn't change a single line in DIRAC**
  - The VM is low priority process and can be paused by the user or BOINC

- It works, but… with some "buts"...
  - So, the registration is closed!



LHCb

Issues encountered

- CVMFS
  - We rely on it, totally.
  - The volunteers can be anywhere, the local cache has to be populated. Squids?
- Jobs length
  - Grid jobs are usually long, and volunteers can just switch off their machines at any time
  - Solution: short jobs using MC “adaptability” capabilities
- Monitoring
  - Who is running this job?

Issues encountered /2

- Accounting
  - How to?
  - Possible solution: know where volunteers are, account accordingly
- Clients
  - You’d be surprised, but the BOINC windows client works better than the linux one! (mac is ok)

Desktop grids

Desktop grids

CERN IT Department

- With BOINC, you can run on desktops or on regular clusters within an institute (e.g. Tier-2 site)
  - Resources in-house, i.e. controlled environment
  - Idle office PCs in general under-used
    - E.g. CERN Desktops outside working hours
  - BOINC clients started as a service on a machine, all running with the same user
    - No need to worry about BOINC credit and user forums, graphics etc
- Easy to setup, good way to get started and use available computing resources!

We would like to have few desktop grid sites!

- Volunteer site in “Universita’ di Ferrara”, in Italy
- Windows PCs, got some issues installing as a service
  - Managed in the end, with passwords automatically generated and encrypted
    - All credits to Michele Ferrari

Conclusions

- Volunteer computing is another paradigm of distributed computing
  - It provides (very) opportunistic resources
  - Not what you were used to do with Grids
- It’s not only about volunteers: in the end, BOINC is a way to easily run MC jobs (almost) everywhere
- Doing it well is doable but not trivial, and you need to dedicate someone to it.
  - LHCb is at the moment, looking for this “one”.

Comments and questions

?