



# Thoughts on upcoming developments in computing

Raja Nandakumar

(With thanks to many LHCb experts)



# Ideally ...

- Move to cloud computing pattern of resource usage everywhere
  - ▣ Crucially the job runs for as long as VO wants
  - ▣ Also the job gets all the resources on the node (multi-core)
- Virtualisation to enable the above easily
  - ▣ Site maintenance much easier
    - ☼ Can move jobs from one node to another
  - ▣ LHCb likely to find it easier to maintain job environments (among other things)
    - ☼ Still need framework for this for user jobs ...

# Clouds

- Mechanism to communicate automatically between the site and VO
  - ▣ Starting and stopping jobs / images on demand
- Trust of the VM by site
  - ▣ CERNVM likely to be widely used by LHCb
    - ☼ CERNVM investigating p2p to reduce network load? (Alice independently in the last CHEP)
  - ▣ Possible independent use for CVMFS for software distribution (move away from software area)
    - ☼ Tests ongoing at PIC
- Problems this will address :
  - ▣ Decouples hardware & OS from job
    - ☼ Some re-write of the batch system?
  - ▣ “No” CPU / wall time limits for jobs
    - ☼ All experiments will love this

# Multi-core jobs

- Likely to come sooner than later
- High efficiency from Belle experience?
  - ▣ Reduce inefficiency during job start / end
- Some WLCG proposals being investigated by LHCb
  - ▣ Shutdowntime, shutdowncommand, hs rating / core, ...
  - ▣ Also making tests on running multi-core jobs
    - ☀ And integrating this into DIRAC
    - ☀ Not yet possible to reserve whole node using grid interfaces
- GaudiParallel already allows us to run multi-threaded jobs
  - ▣ Pilots should also be able to run parallel jobs if needed
    - ☀ In parallel Python threads for example

# Interesting possibility

- Scatter-gather mechanism on a site - level
  - Particle physics very well suited for this
    - ☀ Data consists of many “events”
  - Should work very well for productions
    - ☀ Possibly for “standard” analyses too
  - Already used in online farms for some time now
  - 100 % fault tolerance
    - ☀ Event picked up by different process if original process dies
    - ☀ Scatter / gather process can be restarted if needed
      - ⇒ Only a few minutes lost
  - Devil in implementation details of course
    - ☀ Both in code and in site management
    - ☀ Not clear it will be useful for all VOs