Remarks on Benchmarking

Helge Meinhard, CERN-IT

WLCG Workshop

Lisbon 01-Feb-2016

Precise Benchmark

- Requires full machine (HW!) to be configured as worker node(s), and full control of the workload
  - As many benchmark jobs in parallel as there are job slots
- Good estimate of “minimum” job slot performance
  - Higher performance if code optimised or if hardware not fully used
- Original target precision: 10%; still not too bad a match with reality…
  - Aiming for reproducibility at the 1…1.5% level
  - Applications, machines, benchmarks have evolved – need to react and follow
- Well, does it need to be replaced? Yes, for procurements; experiment requests, pledges and installed capacity; accounting
- More details: Manfred Alef’s talk

Fast Benchmark

- Why – Is HS06 not sufficient? Well… no:
  - Amount of resources increasing where we do not control the entire hardware (HPC, volunteer computing; clouds
  - Needed for job masonry; accounting; procurement of cloud resources
- Less precise and reproducible
  - No control of other jobs' workload on same hardware
- Need regular and fast execution
  - Between payloads, or even concurrently
- Number of candidates
  - CERN used ATLAS KitValidator (KV) in commercial cloud procurements
  - More work needed
- Can we achieve a single solution for WLCG across experiments and workloads?

Accounting

- Acute or latent suspicion about accounting numbers being inaccurate
- Expectations need to be set correctly
- Accounting is surprisingly complex; try and make it as simple as possible
  - Both at the site and at WLCG/EGI level

Storage and Information Transport

- Agreed approach: Machine-job features
  - Deployment is easy (still risks to take long due to chicken-egg situation)

Need for perfection?

- No, we need something that is good enough – not more
- … and we need something that is compatible with mainstream / commodity
  - Cloud providers may or may not be able to do something special for us… if they are, it risks to be expensive