Status of Benchmarking

Helge Meinhard, CERN-IT

WLCG Management Board

14-Jul-2015

Outline

- Purposes of CPU benchmark
- HEP-SPEC06
- New benchmark: motivation, status
- Requirement for fast benchmark
- HEPiX benchmarking plans

Purpose

- Describe pledges and installed capacity
- Provide usage information (accounting)
- Serve as metrics for procurements

HEP-SPEC06 (1) (abbr. HS06)

- Current work-horse for CPU benchmarking
- Established in 2008/2009
  - Proposal by HEPiX working group adopted by WLCG
- Aim: reflect behaviour of typical applications on WLCG compute farms to some 10% accuracy
- Choice:
  - Application: Subset of SPECcpu 2006 integer and floating-point test suites written in C++
  - To be run on OS typical for WLCG compute farms
    - Currently RHEL 6 x86_64 compatible, gcc 4.4
  - Strict choice of compiler flags:
    -O2 -pthread -fPIC -m32
    - Following advice by Architect's Forum at the time

HEP-SPEC06 (2)

- Reason for choice of C++ subset of SPEC cpu 2006:
  Good agreement of relevant low-level CPU counters between these applications and compute farms with LHC load
  - FP over integer instructions
  - Wrong branch predictions
  - L3 cache misses
  - …

HEP-SPEC06 (3)

- Single LHC applications may differ significantly from HEP-SPEC06 behaviour
  - LHCb called for offers for HLT farm nodes using their proper trigger application rather than HS06
- Typical mix, as seen on lxbatch, still scaling with HS06 to the ~10% level
- No immediate problem, but…

New Benchmark – Why?

- SPECcpu 2006 support will run out soon after release of new SPECcpu suite
- Benchmark based on 32-bit applications with low optimisation no longer fully adequate
- HS06 rather more sensitive on speed of memory than typical applications
  - Issue with interpretation of perfmon counters at the time?

New Benchmark – Status

- SPEC known to work on new release of SPECcpu
  - Initially expected for end 2014, now rather for 2016
- Potential alternatives:
  - Application commonly used by all LHC experiments
    - Geant 4 application being tested
    - Advantage over SPEC: licence-free
- Industry standards maintained by third parties much preferred over HEP-specific solutions

Fast Benchmark

- HEP-SPEC06 typically runs for 5…6 hours
- Fast(er) benchmark required for quick-and-dirty normalisation of resources (e. g. when a VM is created)
  - Not suitable for capacity management and procurement
  - Several candidates:
    - Python script provided by LHCb
    - Hammercloud application developed within ATLAS

HEPiX Benchmarking Plans

- Benchmarking working group being revived
  - Co-chairs: Manfred Alef / KIT, Michele Michelotto / INFN
- Potential contributors contacted
  - More welcome – feel free to contact co-chairs
- Studies started around fast benchmarks and Geant4-based application
- Group waiting for next SPEC release

Summary

- HEP-SPEC06 has served WLCG very well over 6…7 years
- No immediate significant problem
- Need to move forward in line with outside community
- Fast, less precise benchmark is required as well
- HEPiX benchmarking working group revived, started work

Questions?