

WLCG HEP-SCORE Deployment Task Force

Meeting on 04 May 2022 at 14:00 h UTC (teleconference)

Notes

Indico event page: <https://indico.cern.ch/event/1104203/>

Welcome, note-taking, notes from previous meeting, matters arising

The minutes from the previous meeting are approved; thanks to Stefano Piano for having provided them. Bernd Panzer will take the notes for this meeting.

Short update on status of workloads and benchmarking runs (Domenico Giordano, Gonzalo Menendez Borge)

Helge mentioned that there is a possible issue with the Gravitational Wave code. The workload is validated but shows only 80% CPU usage, which Domenico did not understand. ALICE proposed a combined GEN-SIM Reconstruction workload, but within a Singularity container the RECO step creates problems (outside of singularity this seems to work fine). Thus, Domenico and ALICE agreed that the Gen-SIM part gets packaged separately. New short notice proposal from ATLAS about another workload (reconstruction of real data) was received. This is being followed-up by Domenico and Walter.

Josh thinks that the 80% CPU usage is probably real. The workload is 4-way threaded and there might be some waiting for threads included.

Stefano said that ALICE is proceeding with the validation, some problems with shared memory have been solved, but a new issue with conflicting libraries in the container has appeared.

Walter mentioned that the issue with the reconstruction-of-real-data workload is the database release. Progress will hopefully be fast.

Helge proposed to have a deadline for workloads to be considered for inclusion into the official HEP-SCORE 2022 release. Domenico needs about 5-6 weeks for the final validation. Thus, all included workloads need to be validated at the latest by end June, which means the selection has to be done by the 15th of May. Even if the specific ALICE and new ATLAS workflows are not ready for mid-May, the work should be still pursued as these are interesting.

There was an update from Gonzalo on the benchmark collection campaign. Most of the sites finished all of the benchmarks, only 3 sites just finishing the last scripts and one site a bit behind. He showed a slide with specific node architectures which do not have a full set of benchmarks.

Analysis of benchmark results (Randy Sobie)

Randy gave a very detailed talk about the validation of benchmarks and their performance relationship to each other (experiment workloads, HEPspec06, SPEC2017). He concluded with three observations: Simulation is the dominant user of the CPU with good agreement; some workloads are more performant with hyperthreading off; some workloads have large variations for different CPUs.

Tomaso was asking about more detailed monitoring during the benchmark runs (like swapping) to better understand the counterintuitive results for some of the benchmark results. Domenico explained that this type of monitoring is available and could be deployed (like perfmon). Domenico explained that the shown plots were normalized versus physical cores, thus the actual 25% increase overall (SMT on versus SMT off) is still there. As shown some single threaded workloads do not profit as much as HS06 from SMT ON, there is some saturation before reaching the full number of SMT-ON jobs. It would be interesting to know whether this less-than-usual scaling of specific experiment workload with SMT-ON is acknowledged by the corresponding experiment experts.

Andrea asked whether we should exclude some workloads which have a large dispersion and also emphasized that extra monitoring to understand some of the dispersion would be needed. Another point he made was that it would be useful to have more plots where the used CPU architecture is disentangled.

Helge asked for an action point where the experiment representatives make sure that some of the seen workload behaviors are real or maybe artefacts of the benchmark infrastructure.

Helge advocates that the benchmark suite should not necessarily contain all available workloads, but rather focus on being representative, especially if the variation of some of them are within 5%. More arguments were exchanged: benchmark should still be as inclusive as possible; benchmark should not take too long to run; having more workloads protects against outrunner.

It would be good to define roughly the length of the total benchmarking exercise, taking into account that this would also run at vendors and experiments would use this also run quickly within pilot jobs.

Andrea reiterated the point of monitoring extra CPU/System related counters to correlated them with the actual benchmarking. Domenico said that this type of monitoring is not done on a regular basis but could be enabled for specific benchmarking exercises. There is though an issues with the required competences available, this should be followed-up.

Action items:

- Continue to validate the ATLAS and ALICE workloads
- Another cross-check whether the experiment workloads representing the behavior in 'real-life' using HS06 as reference
- Domenico to investigate the availability of low-level CPU monitoring expertise in CERN IT (Markus team and procurement team)

Any other business

Helge vented the idea of having a longer meeting or workshop to have more time for intense discussions. September/October would be an appropriate time and possibly co-located with existing events like HEPiX or as

pre-GDB meeting. There seems to be some preference to hold it CERN as a maybe two half days event. More input to this idea should be sent via the mailing list.

Next meeting

First and third Wednesday every month

Next meeting already scheduled for 18 May 2022

Annex: Attendance

Present:

Miltiadis Alexis (CERN)
Tommaso Boccali (INFN Pisa)
Domenico Giordano (CERN)
Michel Jouvin (IJCLab)
Walter Lampl (U Arizona)
Helge Meinhard (CERN, chair)
Gonzalo Menendez Borge (CERN)
Bernd Panzer-Steindel (CERN; notes)
Stefano Piano (INFN Trieste)
Oxana Smirnova (U Lund)
Randall Sobie (U Victoria)
Andrea Valassi (CERN)
Josh Willis (Caltech)
Tony Wong (BNL)

Apologies: Gonzalo Merino (PIC), Andrew Melo (Vanderbilt U),