Helge Meinhard / CERN *11 February 2022*

WLCG HEP-SCORE Deployment Task Force

Meeting on 02 February 2022 at 15:00 h UTC (teleconference)

Notes

Indico event page: https://indico.cern.ch/event/1104193/

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

The minutes from the previous meeting are approved. Thanks to Walter Lampl for preparing them.

Status of benchmark runs (Gonzalo Menendez Borge)

Gonzalo MB presented the slides attached to the agenda, emphasising additional contributions of benchmarking configurations (thanks) and encouraging sites to re-run HS06 in the context of the scripts even if earlier measurements exist, in which case they can be cross-checked.

Helge asked to indicate reasons (other than missing licences) why the SPEC-related measurements are not complete yet, and emphasised the importance of cross-checking results for similar configurations across sites as well as ensuring that popular configurations at our sites are well represented. He also invited sites to propose ARM-based systems, and underlined that we need confirmation that the benchmarked systems are used in the same configuration for running workloads.

Replying to a question by Michel, Gonzalo explained the multiple entries in the configuration table (different disks and SMT configurations). Michel suggesting a better grouping of entries. Domenico added that other views to the dashboard give this information; so far no single CPU model is covered by multiple sites and emphasised the importance of comparing HS06 measurements in the framework of the current scripts with earlier measurements.

Status of workloads (Domenico Giordano)

Domenico explained that progress is being made on ALICE and GW workloads. For ALICE, the conditions database is being integrated with the container as an SQLite database, and the result extraction is being implemented. For GW, failures of the container on large machines are being investigated.

Stefano provided additional details for ALICE, expecting to release the updated container before the end of the current week. Josh added that the version sent to Domenico is taken directly from git and is awaiting a formal release stamp by LIGO, which is expected once the changes are verified. Domenico asked that GW uses the standard workflow of the benchmarking developers; Josh agreed to try, but pointed out that he needs a slightly different environment for his local tests.

Replying to a question by Emmanouil, Domenico explained that the first round of HEP workloads will include the new version (SHERPA-based) of ATLAS gen and the multithreaded ATLAS sim. Walter added that work is going on on creating a database snapshot required for ATLAS reco. Emmanouil pointed out the large memory footprint of the pileup in the Singularity image of SHERPA; it was agreed to follow up offline. *(Note added after the meeting: this was clarified and corrected.)*

Observations from various benchmark runs at GridKA (Matthias Schnepf)

Matthias presented the slides attached to the agenda.

Replying to a question by Helge, Matthias explained that potential explanations for the lower performance of some AMD EPYC systems when using twice the physical cores with respect to 1.5 times include thermal throttling as well as overload of some processor units, for example the branch prediction unit.

Domenico pointed out that what is called HEP-SCORE on the slides is the prototype workload collection (based on Run-2 software) defined previously and not what is being launched now (based on Run-3 software for the LHC experiments). He explained that the fact that HEP-SCORE results are similar to HS06 64-bit for newer processors is by design. The interesting statement is the spread of the ratio between the two (slide 16) rather than the absolute normalisation. The observed difference of 20% for some of the CPU models is being investigated and will be followed up with the new workloads.

Any other business

Next meeting

Already scheduled for 16 February 2022

Annex 1: Attendance

Present:

Domenico Giordano (CERN)
Michel Jouvin (IJCLab)
Walter Lampl (U Arizona)
Helge Meinhard (CERN, chair)
Gonzalo Menendez Borge (CERN)
Gonzalo Merino (PIC)
Bernd Panzer-Steindel (CERN)
Stefano Piano (INFN Trieste)
Matthias Schnepf (KIT)
Oxana Smirnova (U Lund)
Randall Sobie (U Victoria)
Jeff Templon (Nikhef)
Andrea Valassi (CERN)
Emmanouil Vamvakopoulos (IJClab)
Josh Willis (Caltech)
Tony Wong (BNL)
Yan Xiaofei (IHEP)

Annex 2: Request to sites

(Excerpt from mail by Helge to task force dated 31 January 2022)

Note the following homework in particular for all site representatives on the task force (but all are asked to contribute):

1. Coverage of (WLCG etc.) worker node models

Please look at the list of platforms available for benchmark runs (see Gonzalo's presentation: https://indico.cern.ch/event/1104189/contributions/4692790/attachments/2375592/4058024/2022-01-19_TF_Measurement_Collection_Update%40TF.pdf, slide 12) and compare it with your installation for running WLCG etc. workloads. If you run systems not (well) represented in that table, please consider offering sample systems for benchmarks. Systems do not need to be permanently available for benchmarking, but should be made available when requested within reasonable delays; during benchmarking campaigns, no other activity should be running on the same physical machine.

For a reliable benchmark to be proposed, it is important that we understand how it behaves on as broad a variety of worker node configurations in use as possible.

2. Reproducibility across sites

Even if you find worker node configurations in use at your site represented in that table by another site, please consider proposing systems at your site. It is important that we know how well results are reproducible from one site and worker node model to another.

3. Machine configurations

It is important that the benchmark runs use configurations identical to those used for real workloads. For example, SMT settings must be identical, the number of cores used must match, virtualisation parameters (if any) need to be identical. While we try and do our best to obtain (or guess) configurations during benchmark runs, we cannot guarantee that real workloads run in the same configuration. So please provide the information in clear text by sending mail to Gonzalo indicating for each hardware configuration that you propose:

* whether SMT/hyperthreading is enabled

* if it is, how many cores are configured for workloads (not all sites use twice the number of physical cores)

* whether virtualisation is used

* confirmation (or not) that the configuration for benchmarks reflects the one for real workloads.

4. More sites and configurations welcome

Points 1 and 2 above imply that the list of benchmark platforms is by no means closed. We will happily add sites and benchmark platforms as proposed. So don't hesitate... it is not too late to contribute!