WLCG HEP-SCORE Deployment Task Force

Meeting on 21 April 2021 at 14:00 h UTC (teleconference)

Notes

Indico event page: https://indico.cern.ch/event/1030671/

Welcome, note-taking, notes from previous meeting

The minutes from the previous meeting are approved. Thanks to Jeff Templon for preparing them.

Status and plans of DUNE workloads (Andrew McNab)

Using experience from WLCG when possible (CRIC for pledges). Fermilab will be Tier 0 when production analysis starts. Have identified the CI suite tests as the workload. Have worked on proof-of-concept HEP-SCORE container; this is not yet ready for a merge request (would not apply cleanly). Some issues and work-arounds identified (see slide 9). Next will iterate with DUNE software experts to finalize the choices for the workload and a plan as the software evolves in advance of the DUNE experiment mid-decade.

DG: Will implement suggestions (identified on slide 9) pending person-power. As far as weighting, believe should run what you find most meaningful, and reweight once there are comparisons with other.

HM: Agrees that it is premature to worry about relative weights of individual workloads, rather put those workloads singly into containers.

AMcN (reply): Need to discuss how to weight so that it is representative of how each workload reflects usage of the overall computing analysis.

AV: Three points. (1) Agrees that should identify which use most of CPU. Twenty seems like a lot; three would be more reasonable. (2) Does Belle II use .cern.ch in CVMFS (answer: yes). (3) Was the shrink-wrap issue only with data (answer: yes). In which case just script to copy it rather than shrink-wrap.

AMcN: Believes failure to shrink-wrap data may be because this is large-scale data that uses XRootD under the hood.

DG: One work-around that has been used is to put the data into a website, and get it using 'wget'.

AMcN: Will investigate options other than putting the data file into git.

TB: After you have resolved the CVMFS issue, will there be other dependencies that require network connection? Answer: No. To Domenico: Would the wget require this? Answer: No, only done when image is created. We run a test that would reveal whether the container needed IP connectivity. Will try to test and document creation of docker archives.

HM: Can run entirely without connectivity once the image is downloaded? Answer: correct.

Next steps

HM: We have heard now from all of the experiments, which are at different levels of readiness. The software in use today will not be what is used in 5 years. How should we organize ourselves, and on what timeline? Helge believes we establish candidate workloads, identify machines and person power to run the basket of workloads, and then see how we can iterate. Therefore we need to decide how to evolve the workloads further on a variety of machines. That shouldn't close the door on changes, but do not want to repeat the effort when workloads are updated. Therefore need to identify sites that can run workloads, and person power to perform the benchmarks. Came up at management board meeting (what is timeline?) Proposed end of the year for a result, but that would require running the benchmarks over the summer. Would propose to discuss this in the next meeting: sites that can make hardware and person power available. Will ensure CERN is included but it should be broader. Would like a presentation about updates to workloads. Would propose this presentation as soon as possible: next meeting (5 May) or later meeting (but 19 May overlaps with vCHEP). Open for comments.

DG: Given that deadline we would need to include new workloads which would require supporting CentOS7— this requires development. Those that run on SL6 can run today. Proposes benchmark with HEP-SPEC06, and with workloads that are ready, and then later run only the workloads that we are still integrating. Data collection can readily be centralized at CERN, if people enable the switch to ship the results to CERN.

WL: ATLAS not completely ready yet; could provide a snapshot for CC7 on the timescale of 1 or 2 months, but not easy to predict how it would evolve (would need to keep at least one software release on CVMFS for benchmarking purposes only).

TB: Would say that CMS is already there—do not expect additional gain in performance in advance of run 3 makes sense to wait for. That is for CPU; for GPU the situation is less clear (infrastructure is not fully in place, e.g. drivers) but this is also a priority.

HM: Only proposing x86 results over the summer; too soon for GPU.

DG: Agree, also would like to support ARM in the future. We need to evolve the infrastructure to allow this.

HM: Making an effort at CERN to find additional person power for running benchmarks and also continuing development of the framework.

MJ: Discussing with colleagues what IJClab can contribute.

HM: Should not delay progress because we want to port something to ARM or Power. So can use whatever already works, but that should not dictate timescale.

DG: Not sure I agree; we want to build the infrastructure to support these other systems. This will be a consequence of the reimplementation, not an additional cost.

HM: timescale?

DG: Had hoped 2–3 weeks, but other person no longer available; hope in one month. Should we just run HS06, and run the basket later? Or would sites prefer to only run only a minimum number of times?

AV: Apologies could not be involved as hoped.

HM: Should get feedback from sites about preference for running more times, but shorter each time, or fewer times for longer.

HM: May I take this as a basic timeline as supported?

AV: Not all sites are present; how will we reach them? Via management board?

HM: Start by asking for contributions from this group at next call (one slide or none).

HM: Hearing no objection, will tell this group that we welcome updates on workloads (since presentation), and sites on what could be contributed in terms of person power and hardware. Then skip 19 May meeting

Any other business

AV: Question about vectorization. We are not using this very much (by and large). Am curious if we should make assumptions about this (e.g. -march=haswell). We should come back to this at some point.

HM: Reminds me of discussion in 2000 that we should investigate compiler flags, etc. Could consider doing this in running the basket.

WL (in chat): We are building for the oldest machine there is.

AV: Considering building fat binaries for benchmarking, even though it is not how the experiments run.

OS: About compiler flags, this is relevant even for serial code

HM: (replying to WL): There is the tradeoff between increasing performance and narrowing machines where you can run

AV: Agree that other flags are important (e.g. —ffast-math) but the vectorization flags would scale very differently on different architectures.

HM: Should have a dedicated discussion. Could argue that it is not the core mandate of this group, but the experts are probably concentrated in this group. Could discuss internally and then reply to management board.

AV: In the end, we are assigning a number to hardware, so we want to know whether that's included in our benchmarks.

(Minute-taker did not get who said this): If it is not used in production, then do not believe it is reasonable to include it in the benchmarks.

Next meeting

Already scheduled for 05 May

Annex: Attendance

Present:

Manfred Alef (KIT)
Tommaso Boccali (INFN Pisa)
Ian Collier (STFC-RAL)
Alastair Dewhurst (STFC-RAL)
Domenico Giordano (CERN)
Michel Jouvin (IJCLab)
Walter Lampl (U Arizona)
Andrew McNab (U Manchester)
Helge Meinhard (CERN, chair)
Gonzalo Merino (PIC)
Bernd Panzer-Steindel (CERN)
Stefano Piano (INFN Trieste)
Oxana Smirnova (U Lund)
Randall Sobie (U Victoria)
Andrea Valassi (CERN)
Josh Willis (Caltech, notes)
Tony Wong (BNL)

Apologies:

Alastair Dewhurst (STFC-RAL)
Jeff Templon (Nikhef)