

IT R&D Advisory Group

Meeting Minutes

23rd June 2021

Present: Tim Bell, Josh Bendavid, Concezio Bozzi, Simone Campana, João Fernandes, Maria Girone, James Letts, Zach Marshall, Helge Meinhard, Bernd Panzer-Steindel, Stefano Piano, Danilo Piparo, Andrea Sciabà

Also present: Marco Cipriani, Eric Grancher

Next meeting: September 22nd at 10:00 CEST.

1 Welcome, agenda, minutes of previous meeting

There are no comments on the agenda and the minutes of the previous meeting are approved.

2 Summary of the Cloudbank workshop and next steps (João)

The Cloudbank is a service used in the US to facilitate the access to commercial cloud resources for NSF-funded research by taking care of several aspects (project proposals, usage tracking, resource cost optimisation and even training). CERN ran a pilot service for a European implementation of Cloudbank using Strategic Blue (SB) as intermediary between CERN and the cloud providers to negotiate prices and optimise spending. CERN procurement funds were allocated to different projects from various experiments and departments, and they are expected to be fully used up by October. During this pilot, we made considerable progress in understanding pricing models and contracts. The goal of the Cloudbank EU pilot is to assess a model for procurement and use of cloud services across organisational units, to complement on-premises capacity (in the case of CERN). Several steps are needed to move from pilot to production, including a better billing dashboard, addressing data privacy and security, procurement and training. As a service, it would be operated by the IT department and an assessment of the requirements for its deployment and operation is needed. Budget holders would take care of the project prioritisation. A second workshop will be held on October 21st and the pilot itself will run until the end of the year.

Danilo praises the solid work done for the pilot but points out that it would not be a high priority for CMS, and he is unsure about the validity of this type of procurement model. The enthusiastic response from CMS was also motivated by the fact that the pilot involved “free credits” to be spent. Experiments manage a budget for operations that cannot be spent on cloud allocations, as it would not be defensible at the RRB; besides, the Cloudbank model requires to explain use cases and negotiate resources, but IT always provided the resources needed.

João answers that a key aspect to consider is the diversification of funding. There are CERN users not in WLCG (for example in TH or BE) who are interested in cloud resources and the Cloudbank model is complementary to what already exists. He also points out the importance of interacting with Cloud providers in a structured way for what concerns data protection aspects.

Zach adds that experiments would seem to be forced to go through CERN to access Cloud resources, while an experiment may well have a direct arrangement with a Cloud provider. He also warns about mandatory minimums in contracts, which may come as a surprise for users.

João replies that this model only concerns what happens at CERN with a CERN budget: experiments as global collaborations are not affected. Concerning user satisfaction, it was surveyed and found to be very positive.

Concezio mentions that people in LHCb benefited from this exercise, but he shares Danilo's concern on the sustainability of this kind of billing model for the LHC experiments.

João recognises that the use cases deployed via Cloudbank are not a priority for WLCG right now and, as Simone said, the allocations that may come from WLCG would probably be minimal amounts and there would be a prioritisation run via this advisory group.

Bernd says that the Cloudbank pilot proves that procurement of Cloud resources can now be done correctly. The key point is the cost comparison between cloud and on-premise resources. We know that Clouds are not yet cost-effective for CERN, and therefore we decided in this forum to have GPUs in the computer centre rather than in the Cloud. We still miss a clear picture of the costs.

Stefano adds that in any case it would be beneficial if CERN could manage the Cloud resources made available to facilitate access and optimise the costs.

João leaves the meeting at this point.

Simone says that if WLCG decides that procurement via Cloud resources is a viable option, the technical solution to adopt is not as important as having first a prioritisation of the process and a clear understanding about the cost model to take informed decisions. In other words, should some new need for resources arise, it should be first discussed in this group, and the recommendation about how to procure it, either locally or from the Cloud and if so how, should be based on solid information about the costs. *The aspect of costing should be the main focus for the second phase of the Cloudbank pilot.* We also should not forget that other WLCG institutes are willing to provide resources and we should not give the impression that CERN is going to solve everyone's problems.

Helge adds that we should remember that the needs of non-LHC CERN users may be (and are) fulfilled by on-premise CERN resources and they do not necessarily have to be catered for via the Cloud.

A question was raised about the statement in the slides about “Contracts in place and services ready to be used: H2 2022”. It was agreed that this point needs to be followed-up.

3 Update on the high-performance machines for interactive analysis (Bernd, Josh)

Bernd briefly reports that last Friday we kickstarted an investigation whose goal is to understand if and how CERN IT can better support analysis activities at CERN. The idea is to have a final report by the end of the year. We know that several types of resources (Ixbatch/plus, EOS, dedicated machines) are being used for analysis, but today it is not clear where it would be best to further invest.

Josh reports on the recent improvements in the interactive analyses the CMG group are running on the high-performance nodes made available by IT. Recently the focus has been on the analysis framework, by improving the level of parallelism with the help of the ROOT team. Currently a full W analysis run requires about 2 hours, but work on further optimisations is in progress. Other workflows are run with a largely reduced CPU load, which results in them being rather I/O limited and were used to measure average data rates on different storage configurations. The second node provided by IT will be used for other analyses; both nodes will be puppetised to facilitate monitoring and HTCondor submission. Open points include home directories and GPU-enabled workflows, but they are not urgent. The speedup expected by ongoing work on the high-level analysis is at least a factor 2, with a possible additional factor 3-4 with a larger machine, which should arrive in September. CephFS as a remote storage solution deserves more testing, as so far it seems to be a very promising solution.

4 Update on SoC on ARM (Tim)

Tim reports on the progress of the activities to provide full support for ARM. The only supported OS is now CentOS 8, and most of the basic supporting services (repositories, Docker images, CERN packages) are ready. The new HEP-SCORE benchmark framework is being tested using containers. Five Ampere servers with 80-core CPUs and 256 GB of RAM will arrive in summer. The aim is to offer production services, including an interactive facility similar to Ixplus and Gitlab central runners for CI/CD.

Maria asks if the new hardware will replace the obsolescent ThunderX-based hardware, and Tim confirms that this is the case.

5 Experiment round table on priorities on heterogeneous hardware

Concezio would like LHCb to be able to use the ARM nodes for their nightly builds. They are also starting to investigate POWER9 (available at the Marconi 100 supercomputer) and it would be essential to have a build of the LCG software stack for this architecture.

Maria reports that the POWER9 hardware promised by IBM has arrived and it is being installed.

Stefano reports that ALICE is porting code for Run 3 and 4 to POWER9 for the same reasons, and there is ongoing work to offload part of the reconstruction to GPUs (mainly from AMD, but eventually also from NVIDIA). GPUs will also be used for ML-based analyses and a new framework for Run 3 is being developed.

Danilo reports that CMS is already running on Marconi 100 and is regularly building the software for ARM; nothing more is expected for Run 3. For the validation of GPU code, CMS is using GPU resources external to CERN.

6 AOB

The next meeting is scheduled for September 22nd.