

## IT R&D Advisory Group

### Meeting Minutes

22<sup>nd</sup> September 2021

Present: Tim Bell, Josh Bendavid, Eric Bonfillou, Concezio Bozzi, Simone Campana, Alessandro Di Girolamo, Dirk Duellmann, Maria Girone, James Letts, Zach Marshall, Pere Mato, Helge Meinhard, Stefano Piano, Danilo Piparo, Andrea Sciabà

Also present: Marco Cattaneo, Eric Grancher, Andreas Morsch, Ricardo Rocha, Markus Schulz

Next meeting: To be decided

#### **1 Welcome, agenda, minutes of previous meeting and news (Maria)**

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

Maria introduces a discussion on the group mandate and the need to reassess it, starting with looking forward at requests for R&D activities ongoing or expected for the next year and their impact on IT services and resources, in particular with potential for collaborative work. To help this process, she proposes to adopt a tracking table, to be joined to the minutes, which contains information and status about the R&D activities relevant to this group. The maintainer of the table will be Andrea.

Danilo and Pere ask if the table will also contain R&D projects carried out in IT and not only in the experiments and EP-SFT. Maria confirms that this is the case: all activities having an impact on IT resources will be covered. Alessandro stresses the importance of reassessing the mandate of the group before defining the tools for tracking. Dirk comments that we are seeing tighter connections between R&D in IT and in the experiments, and in the future the group might focus on deciding together which activities should have priority, depending on the amount of interest they attract. James expresses the concern that too broad a scope for this meeting could generate an unsustainable amount of work to populate and maintain the tracking table. Maria answers that the table is simply a tool to complement the minutes, and we should have as a goal to understand how to make this meeting more effective. Stefano proposes for the next meeting to discuss a list of R&D activities in IT relevant to the experiments. *Maria concludes by saying that the input received today from the stakeholders will be considered for the proposal of a new group mandate.*

#### **2 ARM status for SoC and build/test (Tim)**

Tim reports that the difficulties in the supply chain distribution due to COVID and some quality issues caused the five Ampere servers ordered in April to be delayed without a guaranteed delivery date, although there is hope for the end of the year. On a more positive note, there are now RPMs for the CERN software and Puppet is ready to instantiate lxplus-arm and Gitlab CI on ARM as soon as the hardware is available.

### **3 CERN IT GPU update (Ricardo)**

Ricardo shows how to request access to GPU resources: the information provided allows to dispatch the requests to the most appropriate services. The GPU usage is steadily growing, and this sometimes requires reshuffling resources to accommodate larger requests. Approximately around half of the usage is for ML applications. GPUs are now fully accessible via lxplus, Gitlab CI and SWAN. Early next year CERN will receive some Nvidia A100 cards, significantly more performant and allowing for physical partitioning, support for which is ready in Kubernetes and will soon be in OpenStack. Profiling for vGPU nodes is now also available. Not all GPU nodes will have vGPUs enabled though, as for some applications there is no reason to do it.

### **4 Other news from IT**

Markus reports that the work on Xcache is continuing, to assess the impact of SSD- and HDD-based caches on analysis workflows. There is an estimate of how many cores an SSD-based Xcache can support, as it will be shown at the next meeting. The plan is now to add more workflows and restart an analysis of the EOS logs to get a cross section of the analysis activities using EOS.

Maria reports some news from openlab. The project on OneAPI with Intel has been extended to 2022. The hardware for the Gauss project has been received and the project with Micron, CMS and protoDUNE has also been extended to 2022. The Nvidia DPUs were allocated to LHCb and will soon be moved to the computer centre to be made more generally available.

### **5 R&D in ALICE (Stefano)**

ALICE is now porting the asynchronous reconstruction on GPUs, to be run on the EPN farm when it is not running synchronous reconstruction (which already runs fully on GPUs). A 2x speedup has already been achieved, but it will be even higher when then porting is completed. Is it planned to develop also digitisation on GPUs. The new O2 simulation and reconstruction code is now fully multi-process, which allows to achieve a much better multi-core efficiency, and significant progress was made in integrating HPC and cloud resources, in the future to be further enhanced by the support of non-x86 architectures. The new analysis framework and the "hyperloop" distributed analysis system achieved a 10x speedup in Grid analysis compared to the Run2 system; two analysis facilities (at GSI and Wigner) will be used for fast workaround

analysis. ALICE is also interested in high performance interactive machines for ML-based analysis.

Maria asks if ALICE is part of the new working group to study the potential of analysis facilities at CERN. Stefano answers that for the moment ALICE is working on developing analyses on the new framework and some request in this direction might come in the next months. Concerning POWER9 and ARM, ALICE is rather manpower-limited and would need support in porting specific libraries and compilers.

## **6 R&D in ATLAS (Alessandro)**

Alessandro describes his ideas on how to best engage the community on the R&D challenges to be faced for HL-LHC. Some areas suffer from a scarcity of manpower, namely distributed computing operations, central services and integration of new tools or new versions. Having people from institutes (CERN IT included) more closely involved with the ATLAS development and operations would be an advantage. Specific areas where this involvement would be particularly welcome are data popularity, ARM and POWER9, databases, ML and software engineering, where there is potential for common development (examples are A Common Tracking Software, Gaudi, Madgraph). However, due to the manpower limitations, it will likely be necessary to concentrate on the most important areas and projects, setting priorities together with CERN IT and EP.

## **7 R&D in CMS (Danilo)**

Danilo identifies a few areas that would benefit from IT central support. POWER9 and ARM are needed to build binaries for HPC allocations and to improve the code, and GPUs for development; however, a scarcity of GPU boxes is slowing down development for the HLT code. In this context, Ricardo's presentation shows that we are going in the right direction. Another important topic is the addition of network usage to the computing model, which would require more functionality in the monitoring. This would also allow to improve network usage. Finally, he proposes a role for this meeting as the place where the experiments can advise IT in the initial phases of R&D, on how to prioritise some ideas over others.

Pere and Dirk express support to this view, *and the tracking table should contain only R&D activities in IT and in the experiments or SFT of common interest and only until they become actual projects; in other words, we should discuss at the strategic rather than the technical level.* Alessandro points out that ATLAS shares with CMS the need for including the network in the computing model. He also proposes to have a first step where we agree on the priority areas for R&D and Danilo concurs with him.

Danilo needs to leave the meeting.

## **8 R&D in LHCb (Concezio)**

Concezio reminds that the GPU-based HLT1 system is now approved for production and Nvidia A5000 cards have been chosen and will be purchased. Still in the online area, there is now a coprocessor testbed, which be used for R&D and will work close to the Event Builder. One project is based on IPU for highly parallel fitting, and another is Retina, for track reconstruction using FPGAs. LHCb shares the need to build and validate their software on ARM and will investigate POWER9 using the Marconi100 resources. The request is for R&D servers in production with access to the whole node, preferably. There are also several ML-based activities benefitting from the GPUs offered by IT and openlab. Work is ongoing to exploit HPC centres (NERSC, BSC, Marconi100), but the effort required is considerable. Finally, he invites contributions for specialised effort to port software stack to non-x86 and optimise it.

Maria asks if the level of support provided by openlab for the ex-Techlab ARM resources is sufficient. Concezio explained that these nodes where used very manually and what is missing is a proper service (not necessarily to be provided by openlab, of course). Tim clarifies that as soon as the ARM hardware is installed, the service level will be the same as for x86 nodes. For POWER9, the only resources available are those from openlab. Simone comments about HPC access by pointing out that some aspects, like date ingress and egress, can easily be addressed in a common way across experiments, but this is not the case e.g., for workload management, each experiment having its own and likely not being interested in using something else. Concezio agrees and adds that looking for commonalities is worth whenever practical and that even when interacting with HPC centers experiments can benefit from presenting themselves in a common way. Maria adds that discussions on this specific aspect are ongoing, and they should become better structured in the future. However, for Alessandro, the critical point to address is the technical level, where ATLAS encounters most of the difficulties.

## **9 R&D in SFT (Pere)**

Pere enumerates various R&D activities in simulation (Geant4 optimisation, fast simulation, GPU-based particle transport), ROOT (to speed up analysis and ML, better multithreading, vectorisation), CernVM-FS (to ease image publication) and experimental technologies (software stack for future experiments).

Maria suggests that the SFT activities on efficient analysis would be a good example for the tracking table, being also related to the IT activity on analysis facilities, with a clear impact on resources. Pere would break them down into more specific subprojects to avoid it being too generic.

Page 5

On the idea of having people from IT or SFT assigned to work for the experiments, Pere comments that this is usually a good thing, and it is already happening (for example in the case of simulation), but this works only if the technology remains reasonably common. Alessandro adds that another very important factor is the trust that experts manage to build in the experiment.

Zack underlines the importance of the SFT R&D being close to the experiments, giving fast simulation as an example, where trying to work together with at least one experiment may produce results and lessons useful for everyone else. Pere concurs and replies that this is exactly what SFT is trying to do.

**10 AoB**

The date of the next meeting was supposed to be November 3<sup>rd</sup>, but it happens to overlap with the LHC review; therefore, an alternative date will be proposed later on. The main topic of the next meeting will be the new group mandate and the input received here, assuming that more discussions will happen before the meeting.

Simone thinks that it emerged today that people consider the current scope of the meeting as too narrow, focusing mostly on resources, while it should be more concerned with R&D discussions and planning in general and asks what the next steps are. Maria would like to bring up in IT the input received today, and the new mandate will need to fit into the general plan on how to reorganise R&D in IT. Simone agrees and adds that what is most important is to discuss the needs and the suggestions brought up today and in the past, rather than the meeting itself.