IT R&D Steering Group

Meeting Minutes

13th April 2018

Present: Helge Meinhard (chair), Andrea Sciabà (scientific secretary), Alberto Di Meglio, Torre Wenaus, Olof Barring, Stefan Roiser, Romain Wartel, Predrag Buncic, Tommaso Boccali, Pere Mato, Ian Bird, Bernd Panzer-Steindel

Next meeting: TBD

1. Introduction and organizational matters (Helge) ([slides](#))

    Helge presents the purpose and the mandate of the group, the proposed guidelines and the program for this meeting. The main points are:

    - Ensure that R&D activities in Openlab and Techlab are aligned with the user needs and the IT department strategy

    - Better understand priorities for experiments, set priorities accordingly, within the available resources and issue recommendations to IT management accordingly

    - Ensure that R&D are properly documented, tracked and their results made public

    - Work in a transparent, open mode and avoid duplication when discussing projects and proposals

    It is agreed that all material will be made available via Indico, at least initially.

    Pere asks if the focus will be limited to Openlab and Techlab or will extend to other R&D activities, e.g. related to HL-LHC. Ian explains that there a clear mandate from the IT management to concentrate on Openlab and Techlab at least initially, and most HL-LHC-related R&D is covered by WLCG. Helge adds that, once this group has successfully addressed the Openlab and Techlab issues, the scope may be extended.

    Tommaso asks whether a list of current activities is available. For Openlab, all details are available in the web site. Ian adds that this information should be made available also for Techlab.

2. Overview of Techlab, current projects and pending requests (Romain) ([slides](#))

    Romain describes the objectives and the modus operandi of the Techlab project. The main points are:

    - Provide an environment with off-the-shelves hardware and as close as possible to standard production to help users in learning how to better utilise current computing architectures and processors

    - Guarantee that all results can be published (e.g. that they are not subject to NDA limitations)

    - Add new hardware according to community feedback and industry trends

- Requests are submitted via a detailed Service Now form asking for a project description and other details.

He also observed a very strong interest on GPUs, much less on low-power architectures.

Tommaso asks if requests can exceed the budget for hardware. Romain explains that there have not been serious shortcomings in hardware or manpower but considerable time is spent in prioritising requests.

Pere asks for a clarification of the difference between Techlab and the rest of IT as a resource provider. Several are identified: a) Techlab resources are not available elsewhere, b) Techlab does not offer software support, c) nodes are not puppetised, hence their management takes more effort, d) requests often come from individuals and without their experiments being aware of it.

Requests that involve a considerable effort (e.g. puppetisation of ARM64 machines) need to be carefully evaluated and come as official experiment requests. This concerns e.g. an LHCb project to port software to ARM64 and run it in a production environment.

Helge reaffirms the concept that our group should form an opinion on the products being evaluated and make recommendations about whether moving them to production.

Tommaso asks what happens if somebody requests a resource not available in Techlab. Romain explains that usually they offer the user something similar or tell her to wait. Olof adds that not everything that is on the market can be bought (e.g. some accelerators require special licences), or in some cases the OS cannot be CentOS. Concerning budget, in the past manpower was more of an issue than hardware investments.

Helge asks if we could ask some of the current activities to give a report. It is agreed that this is appropriate for bigger projects, while smaller, short-term activities can be simply mentioned by Romain.

3. Overview of Openlab, current projects and pending proposals (Alberto) ([slides](#))

Alberto describes the process used in Openlab to define a project. Very shortly: overall objectives are defined e.g. in technical workshops, and proposals from either CERN teams or companies are put forward. After a brainstorming to find the best partners, a programme of work is defined, resources are negotiated and after sorting out the legal aspects, the project starts.

In 2017, companies contributed with funding for about 20 FTE, 250k CHF worth of equipment, technical support, consultancy and training. A short description of the ongoing and planned projects follows; all details are available on the Openlab website.

Helge comments that it would be interesting to see also the manpower cost for CERN, not only for companies. Alberto replies that manpower costs fall into two categories: a direct cost for people directly working on a project, and an indirect cost that comes e.g. from

installation and management of resources. Usually companies contribute to both, but in a few (not very successful) cases, they just provided the hardware.

Pere asks what the best channel is to provide input to the overall strategy. The yearly workshops are the best opportunity for it.

Alberto makes a distinction between "tactical" projects, which are short, low-key and involve just a couple of entities, and "strategic" projects, which must fit in the overall strategy.

4. Open requests around GPUs (Romain) ([slides](#))

Romain explains that requests coming to Techlab are either official or "grassroots". Grassroots requests amount to 80-90% of the total and usually require GPUs, look more "random", do not have any long term planning, involve very preliminary testing activities and typically get a lower priority. He then proceeds with listing all requests from the last few months.

Romain stresses the advantage of introducing GPU virtualisation via Openstack, which will make GPU sharing much easier. Some effort is put into avoiding overlaps with Openlab activities.

Torre proposes that experiments are always notified about requests from their collaborators, not necessarily to veto them but to be aware of them. Tommaso agrees with Torre's proposal. It recognised as important that the experiments and our group will be automatically notified of new requests.

The discussion touches again ARM64 and Ian makes a distinction about initial R&D work, which is not a problem, and providing a production-quality environment for ARM64 resources, which would be a huge investment, requiring to divert effort from other activities. Currently, there is little evidence that ARM64 will become a viable alternative. Bernd suggests that if ARM64 has a strategic value for an experiment, the experiment should clearly say so. Stefan proposes to have the ARM64 R&D work done in LHCb presented here.

It is agreed that it is much more important to focus on GPUs (and accelerators in general) than on ARM64.

In any case, Helge proposes to put ARM64 in the agenda for the next meeting, followed by FPGAs.

Again, on how to get notified of new Techlab and Openlab requests, it is suggested that a summary is generated once per month with a list of tickets (in the case of Techlab) and a similar summary is prepared by Alberto for Openlab.

The date of the next meeting will be fixed in the next days.