



RCS-ICT Technical Committee

The RCS-ICT Technical Committee is an inter-departmental and inter-organizational governance body. Establishes and maintains a roadmap of ICT services and projects jointly engaging the communities in the RCS sector.

Notes of the 8th RCS-ICT Technical Committee held on Friday 6th October 2023

Date: 6th October 2023

Main topic(s): Speeding up MC generators, The Preveessin Data Center, Innovation Roadmap, LXPLUS future scenarios

Agenda (<https://indico.cern.ch/event/1233304/>)

- Speeding up MC generators
- The Preveessin Data Center
- Innovation Roadmap
- LXPLUS future scenarios
- AOB

Attendance

Present: Alberto di Meglio, Andreas Morsch, Andreas Peters, Armin Nairz, Ben Couturier, Ben Jones, Bernd Panzer-Steindel, Elizabeth Sexton-Kennedy, Enrico Gamberini, Eric Grancher, Gavin Mccance, Gerardo Ganis, Gianluca Cerminara, Gloria Corti, Ismael Posada Trobo, Jakub Moscicki, James Robert Letts, Lorenzo Moneta, Marc Dobson, Markus Elsing, Maria Gironi, Massimo Lamanna, Micha Moskovic, Michelangelo Mangano, Oliver Keeble, Pere Mato, Roberto Salerno, Simone Campana, Stefan Roiser, Stefano Piano, Wainer Vandelli, Zach Marshall, Zhechka Toteva Sinanis

Adoption of Meeting Minutes

Zhechka thanked the contributors to the minutes' review. **Minutes were approved.**

Speeding up MC generators (Michelangelo Mangano)

Main Messages

- 20% of computing resources for Event Generators
 - Speeding up (it can bring non-negligible savings)
 - Combination of new algorithms/models and use of advanced computing technologies
 - In addition, need to maximise precision and control of systematics, and maximise event-generation efficiency
- Cooperation of theorists, experimental physicists, and computer scientists
 - IT engagement is critical to envisage the opportunities offered by new heterogeneous computing resources
- Ongoing initiatives
 - Studies by MC authors (GPU, ML), Pilot project in IT to port MG to GPUs, MCnet follow-up supported by LPCC, WS 13-14 November
- Formalising and structuring these efforts has strategic value for the future of HEP, and CERN should play an important role
 - Same priority as given to accelerator and detector R&D
 - Optimal hosting for this activity was discussed
- Difficult to find/recruit people with the right mix of skills (TH, EXP, CS, HPC)
 - Large experiments do have some of them
 - Recognition by the Community and CERN is essential

Discussion

- This project feels more like a software engineering problem to ATLAS offline. The code (of Madgraph) needs to be ported to run on GPUs. The question was raised as to why the IT department should be involved in the project instead of the SFT group.
 - a. Keeping the team working together is crucial.
 - b. The IT department plays an important role in the project since they have access to the hardware and market connections. The project goes far beyond software engineering.
 - c. CERN management decides where the effort should sit and who holds the ownership of the project.
- There is no specific request or projection at the moment. A clearer picture can be drawn when the situation will be understood better after November's workshop. This presentation aims to provide a common perspective.
 - a. There have been efforts invested to port event generators to GPUs. It is important to foresee the model for the future. In the long run, the Event Generator authors should take over the ownership for the code maintenance.
 - b. One of the important goals is to onboard the communities to do things in a common way. It is clear that this is a project that is not going to be exhausted.
 - c. CERN is not the only contributor
- Separation between the common infrastructure design, which will be determined and implemented by the core team, and module development, which can be delegated to students, would be beneficial.
- The needs must be classified and categorised.

- It is necessary for the developers working in this area to be recognized in order to keep their contracts.
 - a. Collaboration between all the contributors is essential.
 - b. One proposal was to encourage publishing the work. The journal to be used should be decided.
- Michelangelo's presentation has been uploaded to the [3rd RCS-ICT Steering Committee agenda](#).

The Preveessin Data Center (Massimo Lamanna)

Main Messages

- One $\frac{1}{3}$ of Préveessin Data Center (PDC) in production at the end of 2023
 - 4MW will be available (comparable to B513)
- The PDC is creating a new opportunities to extend collaboration with the experiments
 - E.g., for hosting HLT computing and disk storage
 - There has been some discussions with the 4 LHC experiments
 - Stringent time-constraints for LHC Run 4!
- Maximising the return of the CERN investment and the resource usage
 - New high-quality infrastructure for the growing experiment systems

Discussion

- In the new PDC utilisation plan, hosting the new High Level Trigger (HLT) should be the easiest target.
 - Having all experiments at the same level will be beneficial.
 - CMS and ATLAS commented that having their HLTs split in two sites is probably adding extra complexity.
 - The idea has been discussed with a couple of experiments; at the end after the heavy duty part, then the system looks quite like a batch system. This should be no problem so far.
 - If the data needs to be sent back from PDC to CMS and then back to CERN, there can be issues with the transfer system.
 - ATLAS will need to change a lot of the online calibration procedures.
 - Both CMS and ATLAS online agreed that if there are some loops, the accounting needs to be changed as well.
 - ALICE needs to move 3.5 TB per second to 1st level processor's farm 1TB from point 2 to preveessin.
 - For ATLAS, a much easier target for hosting compute in the PDC would be the "old" (Run 3) HLT hardware, to be removed during LS2 (early 2026). Installing that hardware in the PDC would allow ATLAS and CERN IT to gain some experience in sharing hardware without the added difficulty of running a critical real-time processing system. It would be excellent if CERN IT could address this possibility first, before we add layers of complexity.

Innovation Roadmap (Alberto Di Meglio)

Main Messages

- The IT Strategy includes the definition of a innovation roadmap with the objectives to support the evolution of CERN programmes and communities needs
- The process started after the IT PoW 2022 with internal brainstorming activities
 - Draft document was opened for internal review in September
 - Proposing now for a general review to all CERN communities
- Document is work in progress
 - After this iteration, it will be further discussed at the next IT PoW in November

Discussion

- The draft roadmap outlines what the innovation section in the IT department plans to do. Community needs must be reflected in the experiments' TDRs and then incorporated into the IT roadmap.
- RCS communities commented that this document is intended for IT colleagues. There are some objectives in the roadmap document that need to be made more understandable for the RCS communities.
- For each objective/project, it should be stated how much effort is required. The description of the project should include this information. Otherwise, everything sounds great, but we need to be able to choose and prioritise
- A roadmap agreed upon by the entire RCS sector is expected to be created by this group. The members of the RCS-ICT technical committees have been working for almost a year in this direction. What is the purpose of a second separate innovation roadmap?
- After agreeing that an activity or project is interesting for the community, the needed resources will need to be evaluated.
- Technical Delivery, Engagement, and Innovation are all part of the IT department, and we want to propose a single common roadmap. The innovation roadmap should focus more on long-term activities.
- There is of course an overlap between the roadmap that the engagement team is creating together with the RCS communities and the innovation roadmap presented today. Many activities in innovation that are perhaps not immediate needs from the RCS community, and many engagement activities do not have an innovation flavour. Both roadmaps must somehow be merged. Discussing the timeframe, the process review, and the cost is important. We should not rush.
- Innovation activities will most probably require additional funding. It is possible to obtain funding from a variety of sources (e.g. 15 MCHF from NextGen). The funding mechanism for these projects must be clear. Drafting a roadmap and collecting comments is an important first step.
- It is necessary to identify the class of each project:
 - 1. we need this but not urgently,
 - 2. we need this urgently,
 - 3. we need this (which is not currently in the roadmap) and it should be added to the roadmap,



- 4. (and we need to be open for) we have not identified a need yet, but this might lead to interesting possibilities in the future.

LXPLUS future scenarios (Gavin McCance, Jakub Moscicki)

Main Messages

- LXPLUS is heavily used by all technical and scientific sectors (2-3k simul. users)
 - General purpose Linux service
 - Exact use-cases from the community are not clear
- Need to define the evolution of it
 - Focusing on the baseline functionality (interactive service, ...)
 - Need to understand people's daily workflows
 - The current setup with global "home" both for batch and user machines makes it easy for users to accidentally overload

Discussion

- LXPLUS and LXBATCH are needed. There was a consensus that different users use LXPLUS differently, but this is a very important infrastructure.
- The statistics show that ~9K distinct users per month use LXPLUS. A survey will be launched to understand their needs for LXPLUS, for example, do they need it to share documents/code with their colleagues. It aims to find out how people use LXPLUS and which features are most important to them.
- Every RCS community should appoint a representative user of the LXPLUS service. As soon as the representatives are identified, the LXPLUS/LXBATCH and File System teams will discuss the survey to be sent to their communities' users. Following the survey results, the IT experts will plan for the pilot with the same representatives.
- Any necessary changes proposed and developed through the pilot program will be rolled out to users after a period of validation and thorough testing.

AOB

ARM runners in GitLab (Ismael)

Discussion

- The announcement that the GitLab ARM Runners will be available to the communities was perceived as a big uptake. We need to be able to assess whether the system is overloaded.
 - CMS Online had a meeting with the systems on chip (SoC) project team the same day and the announcement would be transmitted. CMS will appreciate getting feedback in the following weeks about how well the GitLab ARM Runners infrastructure scales.



Action List

- **2023-08-A1 - Simone/Alberto**: Clarify the overlap between the engagement and the innovation roadmaps with the idea to produce a common roadmap.
- **2023-08-A2 - Gavin/TC chairs**: Create “task force” with one representative from each RCS community (LHC experiments, SME,..) with the goal to develop a survey to be sent to the respective communities in order to find out the exact usages of the LXPLUS service. The lifetime of this task force should not be very long. The TC chairs will put the request to the TC members to appoint somebody from their communities.

Comments/Amendments

- Thanks to Zach Marshall and Stefan Roiser for their valuable contributions and corrections to the notes.