Prospectives LCG-France 2020

Conseil de direction LCG-France

July 19, 2019

version 1.3

1 Introduction

LCG-France coordinates the French computing contributions for the LHC experiments in the context of WLCG. To this effect, MoU were signed between French sites and WLCG. Within financial constraints defined in an agreement over the period 2018-2022, the quantitative goal of LCG-France is to maintain its contributions to LHC computing resources at the level of 8-10%.

The LHC computing grid was very successful in Run 1 and Run 2: despite increasing demands for reconstruction (thanks to the performance of the LHC), simulation and analysis (high precision demands large statistics and refined analysis methods), WLCG delivered data and computing on the timescale required by the experiments. This was only made possible by a constant adaptation and revision of the experiment software and tools and of the WLCG organisation, operations and tools. This evolution will continue during Run 3 but Run 4 (aka HL-LHC) might need a revolution given the challenges of handling the huge datasets of ATLAS and CMS. The success of the LHC computing grid has attracted other HEP experiments (e.g. Belle II, DUNE, ...) and astro/cosmo experiments. This requires further R&D so that grid computing can better accommodate non HEP scientific domain needs.

This contribution to the prospective studies "Calcul, algorithmes et données" mainly focuses on the roadmap for computing at HL-LHC but one has to underline that ALICE and LHCb have committed to extensive overhaul of their computing models for Run 3: huge computing facilities are built on site and for LHCb will produce almost exclusively data under the final analysis format. The LHCb needs for simulation will also very significantly increase.

Part of the computing landscape is better defined in the experiments (we know of a contribution from ATLAS) and technological R&D groups (we know of a contribution from DOMA-FR). We have tried to minimize potential overlap with these contributions. We will be discussing the possible evolution of the organisation, of French computing sites and technology.

This contribution only focuses on national topics and the reader is expected to be familiar with WLCG and the WLCG strategy document.

2 Evolution of the organisation of LCG-France

2.1 Scientific and computing context

Large non-LHC HEP experiments have recently started (e.g. Belle II) or will start relatively soon (e.g. DUNE). Some, like Belle II, already share the LHCOne network with LHC experiments. Most uses the same compute and storage centers with LHC experiments. Other fields, like radio-astronomy are confronted with large datasets and future experiments like SKA even more (LHC-size). As a consequence, in a submission to the European Strategy, the chair and deputy of WLCG proposed a new organisation. It was proposed that the operational part of the current WLCG organisation (GDB, monitoring, accounting and working groups) would be opened to non-LHC experiments. That would allow a better coordination of infrastructures and a re-use of computing tools developed by the LHC community. If validated, such a change of WLCG organisation raise the question of the current LCG-France organisation.

In the past few years, there have been concerns within the LCG-France community about other HEP and astro/cosmo experiments and their computing models. Even within IN2P3, they vary greatly in their computing needs and organisation (Belle II, CTA, DUNE, EUCLID, LSST,...). If a large computing center like CC-IN2P3 can cope with the diversity of needs, Tier 2 administrators would not easily be able to support very different tools and computing requirements. Some level of planning and coordination was suggested. We already have experienced problems in network transfers that were caused by stress tests from a non-LHC experiment we were unaware of.

We know that current radio-astronomy experiments like LOFAR have large datasets (e.g. there is a 10 Gb/s RENATER link from Nancay to GEANT for data recording at Nijmegen). SKA will have huge datasets and plan for a European analysis center that could involve France.

France, and in particular institutes involved in LCG-France, is contributing to several European Union projects like ESCAPE or XDC that are targeting at a wider scientific community. Some of the goals, in particular those in which IN2P3 labs are involved, are related to the R&D efforts currently ongoing for HL-LHC and making sure the selected tools are suited to non-HEP communities.

Historically, WLCG and EGI have very strong connections. WLCG still relies on some of the services and tools from EGI. In the past few years, and more so in the projected evolution, WLCG has been moving away from EGI solutions. Conversely, EGI seems to be more focusing on cloud computing and less on grid computing. In practice though, a large part of the infrastructure is common or

co-hosted. At the national level, the roles and cooperation between LCG-France and France-Grille should be clarified.

This would suggest that a loose coordination of experiment computing experts and R&D effort representatives would be beneficial at IN2P3 level and possibly at a higher level. For networking, a meeting between representatives of the major consumers, CC-IN2P3 and RENATER was organised. We found it very beneficial.

2.2 LCG-France organisation and funding

Over the period 2013-2022, LCG-France has been organised under the conditions of an agreement between the funding agencies (IN2P3 only for the second period), sites (except GRIF-IRFU for the second period) and the LHC experiments. It is felt that such an agreement is very beneficial in both securing funds and defining rules and "dues". We would like to see a similar agreement on a similar basis be negotiated for the period that leads to the HL-LHC startup and beyond. Such an agreement should include funding at least at the current level.

Tier 2 and Tier 3 are also funded from local projects at laboratory, university/school or regional levels. On the period 2013-2016 from which we have collected all the relevant information, this represented more than 300 k€ per year of direct funding and approximately the same amount in electricity cost not paid by IN2P3. It is important to maintain an organisation that facilitates and possibly attracts such funding projects. In the same spirit, there are European Union funded projects that are related to WLCG (either promoting R&D useful to us or promoting solutions from our community for use in other communities) that can help support the effort in our sites.

The ALICE and LHCb collaborations are installing extensive upgrades to the detectors and consequently to their Computing Models for Run 3 while ATLAS and CMS have scheduled massive upgrades for the HL-LHC that will represent a huge step in data size and complexity. Currently, and in the agreement running up to 2022, the share of resources per experiment at CC-IN2P3 have been unchanged since the startup of the LHC and are roughly proportional to the number of French physicists involved in each experiment. Since the computing landscape is changing, this could be a topic of a future agreement.

Since there is great tension in being able to fund the resources needed for HL-LHC, we think it is important to continue the current effort in France to participate in the costing of facilities and experiment workflow as part of DOMA-FR.

3 Evolution of Computing Models and the French sites

There is currently a large R&D effort for HL-LHC computing. It is well known that simple extrapolation of the current models and technologies fall short by a factor 5-10 of the HL-LHC needs. Suggested changes to the analysis data format could alleviate part of the problem. E.g. in CMS the nanoAOD data format has been set in place since 2018, allowing to cope with more than 50% of the user needs for their final analysis format. If such an approach could be extended to the vast majority of analyses in ATLAS and CMS for HL-LHC, it would allow for a major reduction of the disk resources needs. Intensive tape usage ("tape carousel") could reduce the need for data on disk. On the other hand the economic future of this technology is uncertain.

A large R&D on data storage and access is ongoing (DOMA) from which emerged the proposal of Data Lakes in which data would be stored on much fewer sites and remotely accessed (possibly via caches) by analysis and/or simulation centers. We refer to the DOMA-FR prospectives for details and interest of the French groups. Such a model would have a very strong impact on the LCG-France sites, and in particular on the Tier 2s. Some sites actively participate in the DOMA R&D in the goal of becoming a storage site in the context of a Data Lake. Most sites have expressed that they would see little advantage to become "diskless" as they would still have to manage storage for local LHC users and other experiments. A possible exception is LPSC which, due to funding and HR reduction, is considering becoming a "diskless" site. For multiple reasons the SUBATECH site will stop after Run 3.

The evolution of the LCG-France sites in the context of the way data is accessed (remotely or locally, the amount of data moved) has a strong impact on the network evolution to cope with our needs. It is therefore crucial to keep a close collaboration with RENATER and to consolidate the current work on the global infrastructure cost (disk vs tape vs CPU vs network vs performance) making sure we evolve towards a sustainable national infrastructure.

The evolution of the Tier 2 sites is also driven by their ability to participate in local funding projects (CPER, FEDER, LABEX...) and we note that in several instances such projects are funded on a broader scope implying non grid specific technical solutions, like cloud computing or industry adopted storage solutions.

LCG-France is mostly funding and coordinating the use of pledged resources with some exception like the Analysis Facility (AF) at CC-IN2P3. For analysis, physicists are relying on the AF or local resources at their institutions (on- or off-grid) or at CERN. The demand for resources is growing and evolving (eg GPU, see below). Should LCG-France evolve to also coordinate or advise on analysis resources ? To reach a conclusion, discussions with sites, physicists and experiments would be needed.

4 Other technologies

4.1 GPU

The usage of GPU in scientific computing has exploded in recent years. Indeed, GPU provide large computing power for tasks that involve moderate amount of data flow, little branching and simple, essentially parallel, algorithms. As a consequence, GPU acceleration of full detector simulation or traditional event reconstruction remain modest. Work is ongoing to be able to more seamlessly integrate GPU-based algorithms in frameworks, like CMSSW. On one hand, GPU are used in event building as part of fast reconstruction at trigger level (we consider these applications beyond the scope of WLCG) and in part of the analysis workflow (ML in particular). The demand for GPU resource is steadily increasing and LCG-France as an infrastructure provider should investigate solutions, in particular understand if access should be part of the distributed computing model (e.g. GPU "grid" queues) or in dedicated off-grid T3 facilities.

On the other hand, "GPU queues" already exist and provide access for people that don't have GPU at their institute or in regional/national HEP datacenters. The usage can grow but user will be faced with the inevitable heterogeneity of the resources. Indeed, it is documented that performances depend not only on the hardware but also on firmware, drivers and host OS type and settings. These arguments would favor dedicated off-grid facilities like the GPU farm at CC-IN2P3 unless there is a strong push from the LHC experiments towards the grid solution.

4.2 HPC

The experiments already have exploited HPC resources either opportunistically or as part of official CPU allocations. In France, ATLAS has successfully used HPC resources at IDRIS as part of a "proof-of-concept" project. In Nantes, the CPU resources from a CPER have been bought as part of an HPC machine. In Strasbourg, tests are planned in 2019 to use the university HPC center to produce Monte Carlo events for CMS. In the US, there is a strong incitement to use the national HPC resources for HEP workflows, so we can expect that experiments will put significant effort to efficiently run a good fraction of their workflows on multicore / multihost resources. It is not obvious that these workflows would be able to make an efficient use of GPU resources that are an essential part of future machines, including the "Jean Zay" machine at IDRIS. From the experience of the ATLAS collaboration, we think that using HPC resources is possible but in the medium term should focus on CPU. In addition, to be beneficial to LHC computing, there should be a minimum level of allocation throughout the entire year negotiated with the HPC site with possible peak allocations though more traditional HPC calls (DARI).

5 Developments

Currently, the French sites are operated mostly using grid specific solutions. Our site administrators are very experienced but still that makes it difficult to attract or recruit new people. In addition, we are not involved in the development of these solutions so have little impact on their evolution. We should discuss as a community if there are projects were we would like to have a stronger involvement.

France has a larger involvement in higher level grid middleware or experiment software developments, to cite a few: DIRAC, the Operation Portal, AMI... There are other tools like rucio (data management) which seems to gain a wide adoption and has adopted an open development model so specific contributions could be planned. More broadly, the DOMA R&D activities have lots of opportunities for contributions, even beyond IN2P3 (e.g. INRIA or other computing researchers). We refer to the DOMA-FR contribution for details.

It is to be noted that the WLCG collaboration is currently working on easing the day-to-day work of site administrators (e.g. by reducing the number of protocols or going towards industry-like ones that can be shared) so that they can get involved in R&D or other activities . This would benefit to the whole community and may attract new people.

6 Skills, training, recruiting

In LCG-France we have network of skilled collaborators with long experience in grid computing. In between nowadays and the early HL-LHC operation, a number of Tier 2 site administrators might retire. From a recent survey, operations at most sites will continue if recruitment at a reasonable level is foreseen, with enough overlap time too allow passing on the knowledge and experience.

As alluded above, there is tendency of shifting technologies to those already widely used in the computing or academic world, this will have to be followed closely to define the profile of skills we need to operate Tier 2 sites during HL-LHC operations.

There are currently a number of grid middleware developments we are involved in and some that we could get involved in to our benefit[1].

Some of the development and administration tasks could be performed by people already in IN2P3/CNRS (they could come through FSEP or NOEMI). If they are to work with the HEP specific solutions, we need to train them. If the community switch to industry solutions, we may more easily find people with the right skills in CNRS, or attract young people on the market, but we would need to train the current site administrators.

[1] We are aware of the urgent need for software development in the experiments or as common projects but consider this topic is not in the scope of LCG-France.