Report from the ICAC

Ian Bird (chair), Peter Clarke (remote), Suchandra Dutta (remote), Peter Elmer, Eric Lancon, Michel Jouvin (remote)
15th March 2019

Introduction

The ICAC met at Fermilab on 14th and 15th March 2019, with several members attending by video. The agenda for the meeting presentations and discussions is here: https://indico.fnal.gov/event/20100/. The committee saw presentations from the Fermilab CIO and from teams in the Core and Scientific computing departments, and we thank the Fermilab management for their hospitality and very much appreciate their openness.

General Comments

Findings

The organisational structure of computing sits under the Office of the CIO, which oversees both the Core Computing and Scientific Computing (SCD) Divisions. This organisation has evolved to support the needs of the Fermilab experimental physics community, over several years. As the lab is now entering a new phase of becoming a host lab for an international collaboration and community in a more international context, we were asked to consider whether this structure is still appropriate and adapted to the needs.

Fermilab is looking at developing processes for international project sharing arrangements, addressing partner expectations, and organising national commitments. This also includes accelerating progress towards a more fully-integrated identity federation with CERN in the first instance, for data handling for proto-DUNE, and also needed for the design work for DUNE.

The Fermilab SCD facility has served the local experiments and the Tier 1 for CMS performed well for many years. It is now at the point of evolution in a changing landscape of how resources are provided, and at the point of becoming “host lab” for the international DUNE collaboration. This is a good opportunity to review how the facility is operated and how services are provided. In an era of constrained funding it is important to look carefully at some past assumptions and to understand how best to adapt for the future.

The funding for computing is not ring fenced, but is part of the operations funding also used for detector development and operations. Computing is often squeezed by the other demands on the budget.
DUNE is a major strategic direction for the future of Fermilab. It would greatly help planning for computing resources at Fermilab, if a DUNE computing model could be described, that

indicates the expectations of requirements at Fermilab as the host lab. This is essential in order for the SCD to be able to plan for service provision, as well as indicate where any R&D may be required.

The facility is clearly well managed, but we identified several areas where potential changes would be beneficial. The lab exists in the DOE laboratory ecosystem and is subject to some external constraints. One of these is clearly cyber security. Here the threat landscape continues to evolve and the Fermilab team has been proactive in addressing this. We recognise their expertise and high level of competence in this and fully support the direction and plans that they have in this area. One complication may be the renewed concerns over managing foreign national access to computing resources. This is something that should be addressed by the science labs acting together to work with DOE to manage the risks, while maintaining the open scientific environment that is essential for international scientific collaborations.

Comments

The needs of the international community should be evaluated, and the computing management should assess which services may be needed to support DUNE and its collaborators in the role of a host lab. Once that assessment has been made, it may be natural to evolve the organisation of core and scientific computing to better support the needs of the extended community. It is clear that Fermilab should work with the DUNE collaboration and its international partners to develop the appropriate portfolio of services required of a host lab; these may include collaborative tools, licensing, and so on. The host lab has an important function, and will provide essential central services, collaborative tools, as well as acting as a backstop for problems elsewhere in the community,

The pressure on, and lack of, computing funding is unsustainable as demands for computing continually grow. In fact in 2018 there was essentially no funding for facilities growth or replacement. The exception to this is the CMS funding which is distinct. It will be essential to establish a distinct budget line for DUNE computing, to be able to plan and manage the Fermilab facilities for DUNE in the same way as is done for CMS, to ensure the Fermilab commitment to DUNE computing can be delivered.

It is also still the case that there are opportunities within SCD to reduce the number of experiment specific solutions and support. With constrained budgets it is essential that commonality should be exploited as far as possible, through shared clusters, share software infrastructure where appropriate, integrating teams performing similar work and so on. This is happening through the use of the FermiCloud, and is to be encouraged.

We did not see a full budget breakdown showing the costs, or a budget plan for the future, particularly to know if there are major expenses foreseen. This would be very useful to see in the next review. In addition, the proposed establishment of a resource scrutiny group, to help set priorities for funding is encouraged.

Recommendations

- Establish a DUNE computing project budget at Fermilab, distinct from the other computing budgets.
- Encourage the DUNE computing management to produce a computing model strawman that can be used by Fermilab SCD to plan for the services and resources that may be required. It is understood that this can evolve, but it is important to have a guideline.
- Set up a Resource Scrutiny Group to advise on resource priorities and appropriateness. This could be an evolution of the portfolio management team. It should advise on all experiments including CMS and DUNE.
- Engage with the CERN SSO team to understand how more rapid progress can be made on integrating the identity management to allow federated ID use between CERN and Fermilab, important for DUNE design work. As a side note, the imminent WLCG choice of identify management system to replace x509 should be noted.

HPC and Exascale

Findings

DOE is making a significant investment in flagship/leadership HPC machines as part of the US government exascale program. The first exascale machine is expected in 2021 at Argonne. Exascale machines will represent a huge computing resource during the HL-LHC/DUNE era.

Despite HEP workloads having difficulties to use these machines efficiently, Fermilab is one of the active parties in trying to benefit from the possible HPC access. NERSC (Berkeley) and ACLF (Argonne) have been the two main HPC facilities used by Fermilab, with good success in competitive awards from Argonne, in particular for CMS. Fermilab plans to make these resources part of the USCMS T1 pledges in the future.

Fermilab plans to make HPC resources an integral part of the compute resources delivered to experiments supported by Fermilab, in addition to, or replacement for, on premise resources. HEPCloud is the key component (see section on Facility) to make this happen transparently for the experiments.

Container technologies, in particular Singularity, are key to delivering the environment required by HEP applications on HPC machines, whose standard environment is very different. Other key issues include the need for external connectivity by most HEP applications, delivery of software, and performant data access for data processing jobs.

Comments

Fermilab efforts to allow experiments to benefit from HPC resources is seen as very positive, in particular for CMS and DUNE where the computing needs will exceed on-premise capacities. Fermilab could indeed set a standard by publishing a usability requirements

document toward the HPC sites to clarify what they need to provide to be more readily useable. Since the meeting : it is suggested that FNAL work with WLCG who are also working on such a document.

However, we failed to see a holistic strategy for the role of these HPC resources compared to the conventional on-premise resources operated by Fermilab and international partners, or other kinds of external resources that may be available (e.g. clouds). If it is agreed that HPC resources have a role to play in the global HEP computing landscape, we think that a proper strategy is needed to define how they fit in the HEP computing ecosystem, taking into account their limitations in the HEP context (resource availability, inappropriate workloads…)

Recommendations

- Pursue the technical work around integration of HPC resources in the global resources offered to experiments
- Develop a strategy to clarify the expectation on the long-term role of HPC resources for experiments supported by Fermilab

Facility Plans

Findings

Fermilab currently hosts 3 distinct components related to scientific computing: CMS, Lattice QCD and another called Public. They rely on core services (network, security, authentication…) common to the whole lab. CMS and Public are based on the same kind of architecture despite being separate clusters. They do not share most of their storage solutions (dCache/Enstore is the exception).

Lattice QCD relies on a different hardware/environment than CMS and Public and was not discussed during the review, even if it is partly managed by the same people. It would be interesting to understand if the LQCD requirement would easily be satisfied by remote HPC access - HPC generally being used in this sector.

The CMS and Public resources, hosted on premise, are only a fraction of the resources required by experiments supported by Fermilab. The need to enable seamless access to external resources, such as HPC and commercial clouds, seen as an extension of the local resources, has led to the development of a portal called HEPCloud,. It includes access to OSG resources, to HPC centers through competitive awards, or access to commercial clouds either with a pay-per-use access or using grants. At the core of HEPCloud is a Decision Engine that transparently matches the user workload to a one the available resources, based on the workload description.

In 2019, a fraction of the the computing cycles were delivered to the experiments from non local resources. There is a plan to pledge some of them as part of the USCMS T1 in the future, including for elastic peak needs.

A number of the services needed for the 3 facility components are increasingly being jointly operated, however some duplications may still exist within the current structure. It would be desirable to remove the separate systems and services (with individual configurations) that experiments at Fermilab currently assume. This is true both for resources hosted locally and those remotely accessed using HEPCloud. The goal is to reduce the duplication of efforts, promote more common solutions and have more flexibility/elasticity in resource provisioning.

The existing dCache+Enstore storage solution is complemented with different solutions for CMS and Public for high performance storage. NAS appliances are currently used for Public, but they are very expensive and difficult, if not impossible, to maintain and evolve. For CMS, EOS has been adopted (developed by CERN), which is both highly scalable and addresses needs for high performance using commodity hardware.

For archive storage, no obvious alternative to tape is apparent but, as many sites rely on tapes, it creates a dependency on an uncertain market due to the very small number of hardware suppliers (drives rely on a single manufacture, media on only 2 producers). However this is not a problem unique to FNAL.

The main concern is support for Enstore, the HSM software developed at Fermilab and used by dCache to interface to the tape system (it is also used by PIC in Barcelona and a site in Russia): development/maintenance is reliant on a single person close to retirement. It is planned to evaluate CERN CTA as a possible replacement in order to collaborate on a solution with a broader support (major project at CERN).

The plan for Scientific Linux is to move away from a customised Linux version (SL) and to move to Centos for Scientific Computing as part of the migration to version 8 (CentOS 8). SL has served the community well for many years, but the original need for a specific version is now no longer there, and it is clear that CentOS serves as well. This will remove the need for ongoing maintenance effort.

The Public component of the facility is based on aging hardware, with no continuous renewal process taking place in recent years. By 2021, about 70% of the hardware (CPU & storage) will be out of warranty. A new cluster is going to be delivered primarily for the US QCD community and will also be made available to the other communities (Institutional Cluster model).

A Local Operations Review was done recently and provided several recommendations.

A new DOE rule regulating access from foreign citizens to computing resources has been issued.

Comments

Many problems are already well identified by the people in charge of computing at Fermilab. Several actions already in progress to address them, in particular to unify the separate compute components where possible , and to reduce dependency on Fermilab-specific solutions.

Common technical choices and operational models for the three components may lead to increased efficiency and savings, and should be encouraged. Whilst we understand the practical difficulties in reaching convergence starting now, we feel there should be a plan to do so over some period of time. The implementation of the Institutional Cluster model is a good step in this direction and can be seen as part of a solution for optimisation of resource usage across projects.

This work should also look at potential duplication of effort within the current organisational structure. Such areas should identified and addressed where reasonable.

No clear plan for hardware replacement and investments for the next years has been presented.

Sustainability of the in-house tape software system is questionable.

The HEPCloud service is a very good approach to make possible a flexible strategy of resource provisioning based on budget constraints and opportunities like possible access to HPC resources, competitiveness of commercial clouds or various grants giving access to non-local resources (academic or commercial). In particular, the Decision Engine is an important component which is a distinctive HEPCloud feature compared to job submission frameworks used by experiments.

The storage solutions for the Public cluster require specific attention. Solutions chosen several years ago have proven not to be sustainable, either technically or financially. As the current NAS storage is too small and cannot evolve, users find workarounds based on inefficient use of other resources such as tapes, which in turn puts other part of the infrastructure under pressure. There is an urgent need to investigate whether the future Public storage architecture can be based on that used by CMS (EOS) or to evaluate a mainstream technology (e.g. Ceph) and assess whether it could be a sustainable basis for future storage services. The lack of funding from the projects using the Public resource makes any change more challenging.

Fermilab remains as the main contributor/developer of Scientific Linux (SL) as the significant part of the community moved to CentOS when version 7 was released and as CERN decided to join the CentOS community. Despite SL having been a significant contribution of Fermilab to the community, a specific distribution is no longer necessary, and the plan is to migrate Fermilab use to CentOS as version 8 is introduced.

Recommendations from the Local Operations Review seems appropriate, in particular the need to have large projects outlining their computing models and methods used for resource estimates, and the need for well defined policies for CPU performance and storage.

Recommendations

- The HEPCloud approach may be of interest outside Fermilab and better connections with the community at large should be developed to ensure that it benefits organizations with similar needs.
- A roadmap should be established to address the storage architecture issues for non-CMS users. Acquisition and operation costs of the solutions to be established should be clearly and carefully evaluated.
- A plan to harmonise the three separate components of the facility should be created, to avoid unnecessary duplication of both staff effort and hardware solutions, recognizing the practical difficulties of achieving this quickly.
- A plan for hardware refresh over the next years should be developed, including a cost analysis for using cloud providers. Future investments in computing should be made along the line of the Institutional Cluster model.
- A plan for addressing a long term solution for the tape management system should be developed in cooperation with partner labs. Planned evaluation of CERN CTA is strongly encouraged.
- Mechanisms for providing seamless access to computing resources through federated identity should be developed in collaboration with DOE and the HEP community.

Software and Computing R&D

Findings

Fermilab SCD provides core development and support effort for several mission-critical software areas, including generators (Pythia, Genie), simulation (the GeantV upgrade), two event processing frameworks and LArSoft (a toolkit used by a number of the IF experiments). Fermilab also contributes a key developer (P. Canal) to the ROOT project. His efforts focus on ROOT I/O, which is used essentially by the whole HEP community.

Fermilab SCD also supports a range of software and computing R&D on a very large number of topics. The presentation on R&D in this meeting was the first time that all of the current ongoing R&D projects were summarized in a single place by the current management.

The software R&D efforts are funded from a wide variety of sources, including LDRD, the USCMS Operations program, DOE CompHEP, DOE ASCR (e.g. SciDAC), ECP, HEP-CCE, etc. The lab was particularly successful in the (latest) SciDac-4 round, with 3 of the 5 funded projects being led by Fermilab SCD.

Connections between the R&D projects and their eventual users were sometimes lacking, e.g. the “HEP Data Analytics on HPC” SciDAC project is described as relevant to LHC, but was unknown to US-CMS.

There are clear examples where initial exploratory R&D efforts led or contributed to larger subsequent projects. For example the NoSQL LDRD contributed to the current ongoing (Coffea) analysis systems R&D effort.

Some of the structure of the software projects and efforts dates from (and is in part a result of) the historical split of CMS into a separate department within Fermilab. The previous management of SCD made the step to begin to integrate CMS software/computing with the rest, however some amount of stovepiping remains.

Most of the R&D efforts primarily fund Fermilab staff, although there are a few postdocs (for them funding is typically split 50%/50% with the research program)

Comments

The software and R&D efforts are impressive but it was difficult to identify an overall strategy. Some of the projects seemed partly the result of funding opportunities rather than a clearly prioritized effort.

The historical fork of the CMSSW event processing framework to build the art stack used by other experiments supported by Fermilab resulted in the independent development of two different stacks despite having a common root and need to deliver largely the same features. Different technologies seem to have been adopted by the two teams to solve the same problems. We recognize that the cost to fully reintegrate the two needs to be evaluated against the benefits at this point and that multiple strategies might be possible as Frameworks evolve in the coming years. However, the strategy being pursued was not clear and should be reviewed.

For several software projects, in particular those related to frameworks, there is a will to see a larger adoption and to work with other experiments but it was not clear what are the concrete steps currently underway or planned to make this happen. We had the impression that Fermilab would like to see its products adopted by the community, but again the strategy to achieve that was not clear. It was not clear that SCD was ready to enter collaborative work where Fermilab will bring its contributions but also adopt developments made by others.

Recommendations

- From the work done to summarize all the ongoing R&D projects, produce a software R&D strategy with some clearly identified topics that will drive the future search for funding and recruitments in this area
- Start a specific review on the two main frameworks developed at Fermilab with the objective to produce a vision and a roadmap for making the efforts around them more and more common and ultimately, if possible, merge them. This process should be

well connected to the wider discussions in the community, in the HSF in particular, to reduce the number of frameworks to maintain and build them of as many common building blocks as possible.
- Foster the participation of software developers at Fermilab in the various working groups of the community, in particular under the umbrella of the HSF, to ensure that developments happening at Fermilab are well known in the wider community and to allow Fermilab to benefit more from what is happening outside.

Summary of Recommendations

1. Create a resources scrutiny group to review requests for computing resources and set priorities for allocations of resources between the experiments.
2. Computing funding for non-CMS resources is not ring-fenced and is part of the detector and operations funding, thus gets low priority. This results in years where no resources can be acquired, or old systems replaced, despite demands for computing resources continuing to grow. Consider how a funding line for computing could be separated to ensure a manageable budget. The consequence of not doing so will be a gradual deterioration of services and equipment.
3. A separate funding line for DUNE (as for CMS) would be useful in order to plan the resource profile appropriately. DUNE computing funding will need to be part of a long term plan and not subject to squeezing by other competing demands.
4. Look at ways to speed up adoption of federated identity use as a building block of collaborative services, particularly needed for DUNE.
5. Draft a high level plan for the strategy of use of HPC resources. What are the main goals of the work in this area? What are the highest priority developments to enable success? The close relationship with ANL could be useful in setting out this plan, and perhaps a more explicit common project with ANL could be envisaged.
6. DUNE needs a strong computing collaboration with visible management. This was found to be very important for the LHC experiments in managing a global infrastructure and having a long term voice and plan. We recommend working with the DUNE computing management to encourage putting in place a clear management structure to interact with Fermilab and their other collaborating computing sites.
7. DUNE should be encouraged to draft a computing model, in order that Fermilab (and other sites) can plan their facilities. A draft plan will highlight the areas that need R&D or testing. Such a draft should be produced this year to enable Fermilab management to plan their services and organisation.
8. Fermilab should have a plan for how it becomes an international laboratory for DUNE, what collaborative tools will be provided, etc. The plan should clarify the responsibilities of Fermilab as a host lab, and as part of the computing model.
9. The future storage strategy requires particular attention. In particular, a vision and a roadmap is needed to address the needs in the Public cluster and a plan should be elaborated to address concerns over the sustainability of Enstore, ipossibly by adopting a solution with greater support in the community.

10. A plan to harmonise the three separate components of the facility should be created, to avoid unnecessary duplication of both staff effort and hardware solutions, recognizing the practical difficulties of achieving this quickly.
11. A big picture strategy for the software R&D should be made, in order to understand how the (many) various projects fit into the overall strategy of SCD in answering its challenges. In particular such a plan can be used to ensure that funding opportunities are actually focussed on priorities. The plan should benefit from leveraging work that is happening in the field outside of Fermilab, for example in the HSF, and projects such as IRIS-HEP.
12. Within SCD we recommend that CMS and other projects should be less stovepiped. This is a source of duplication of effort and inefficiency. This must be avoided for DUNE. Facilities and services should be as far as possible common across supported experiments, focussing on function rather than specific requested solutions. We encourage the computing management to continue to re-evaluate the organisational structures in the light of constrained resources and with an eye to the evolving needs of the lab and the experiments.
13. It is essential to have an open, collaborative scientific environment, based on federated identities and trust with other national and international partners. For this reason, ongoing separation of business and open scientific environments is important and must be actively continued.
14. We suggest to investigate having a coherent programme of summer students (or graduate students?) as a potential source of new recruits. Potentially in partnerships with universities, particularly local ones such as University of Chicago where many links exist. Having students and R&D illustrates some leadership capabilities.