FCRSG 2020 Review Report

May 4-5, 2020

Committee: Lothar Bauerdick, Pushpa Bhat (Chair), Brian Bockelman, Taylor Childers, Ian Fisk, Maria Girone

Overview

The Fermilab Computing Resource Scrutiny Group (FCRSG), formerly SCPMT, meeting for 2020 was held May 4-5, 2020. The goal of the review was to evaluate the ability of Fermilab Scientific Computing Division (SCD) to respond to computing resource requests from Fermilab experiments and provide feedback/guidance to SCD and the experiments. The review also provides a mechanism to ensure that the SCD efforts are aligned with the Lab's priorities.

The Charge to the committee, which can be broadly categorized as (1) to evaluate challenges for SCD in its operations, and (2) to evaluate resource requests from experiments/projects, is appended along with the committee responses provided at the Closeout. The committee heard presentations from the CIO, the SCD Head, and the management team as well as from individual experiments regarding scientific computing support for the intensity frontier experimental program at Fermilab, cosmic frontier experiments, and the CMS experiment. Conventional computing resources and requests, HPC resources, and scientific computing service requests were covered.

The committee commends the SCD on its management of its resources and the excellent support and service it is providing to the experiments and projects. The committee also commends recent measures implemented by the SCD to in order to improve the efficiency of resource usage and for continuing to develop and provide tools to users for better job management and end-user experience. However, with the continued constraints on budgets in conjunction with growing demands from current and experiments to be coming online, several challenges and risks need to be addressed, which were discussed at the meeting.

The committee's findings, comments, answers to the questions from the Charge and recommendations are presented below.

Meeting agenda and presentations can be found at
http://indico.fnal.gov/event/23475/

Findings

General/SCD

This was the first meeting of the group under the new name "Fermilab Computing Resources Scrutiny Group". The name change signifies the intention of the SCD to adopt international norms and procedures for the review, to closely follow the CERN model, recognizing that the experiments and projects are becoming more international in nature.

SCD recognizes the need to adapt to the changing computing landscape while preparing to support the upcoming major experiments. However, SCD faces challenges as the operations budget has continued to decline in recent years while the requests for SCD resources and demands for support keep growing.

After the presentation of the Charge and responses to the committee's previous recommendations, the committee heard a presentation on the CERN Computing Resource Scrutiny Group (CRSG) and a series of talks on facilities and services given by the SCD leadership team.

There was discussion with division management as to whether the level of scrutiny as with the CRSG is needed here. The committee also asked the SCD management regarding the anticipated relationship of this committee with the international advisory committee being set up and was told that the advisory committee is expected to provide advice on strategic vision and facilities for Fermilab Computing.

Presentations:

- Welcome, Charge, and Follow up from Last Year's Recommendations
- How CMS Reports to the CRSG Board at CERN

Compute Services

Fermilab computing services deliver significant on-site CPU resources (210 million CPU hours/year). It additionally has test nodes in the Wilson cluster, which include hosts whose architectures match various ASCR resources. These nodes provide the ability to compile and test code on single accelerators - but appear insufficient for testing scaling.

Beyond the on-site resources (and excluding CMS usage), the computing services team was able to deliver approximately an additional 10% of CPU resources through the OSG and 0.5% using HEPCloud. The committee was surprised to find that the HEPCloud did not have the significant profile that was anticipated previously.

Presentation: Compute Services

Facility Services

The Scientific Computing Facilities (SCF) Department provides support and services to manage equipment in data centers and elsewhere. The SCF manages system administration, DAQ computing support, server & disk storage purchases and support including for container and services infrastructure and Scientific Linux. SCF also manages the interactive nodes on the General Physics Computing Facility (GPCF).

There are some requests from experiments which SCF finds challenging to support, such as university-purchased analysis servers, and support for native use of Docker containers.

Presentation: Facility Services

Storage Services

Storage services are managed by the Scientific Data Services (SDS) department, which provides a full suite of services through the data lifecycle - from home directories to disk buffers to tape archives to data management development.

Out of the 17.5 PB public dCache disk, 10 PB will be out of warranty by end of 2020 and the purchases are not keeping up. SDS would like to look for alternatives to dCache and Enstore. The majority of the effort and material costs go into tape archive while the least effort (<1 FTE) goes to storage development. No effort was identified for storage R&D.

Presentation: Storage Services

R&D Efforts

In the future HL-LHC and DUNE era, Fermilab expects a significant increase in requirements for computational resources. To avoid having to dramatically expand on-site CPU capacity, SCD is working to leverage external computational resources (for example at DOE Leadership Computing Facilities) while maintaining data repositories on-campus. While SCD identified a future strategy of maintaining data repositories and data management services on premise, the storage services presentation identified no R&D efforts toward this goal.

R&D efforts to date have focused on porting applications to run efficiently on the new computing architectures (including through projects receiving CCE funding) and on extending the Fermilab infrastructure to overlay other resources (HEPCloud). For HEPCloud, no plan was presented to allow access to resources beyond NERSC, while the division has a goal for 2020 to utilize Theta at Argonne.

Presentation: R&D Efforts

Experiment/Project-specific

DUNE:

DUNE collaboration has set up a Computing Consortium with an elaborate organization structure covering various aspects of data and software management and production. A TDR on software design and computing is also in preparation. SCD management stated a desire to grow the international aspect of the computing contributions for DUNE.

The collaboration plans are to keep raw data indefinitely and keep the latest two versions of reconstructed and simulated data. There are no big jumps in CPU requests by DUNE through 2022 relative to last year. dCache and Tape storage, however, are expected to double or more than double by 2022.

Mu2e:

Mu2e has a modest request for computing resources (summing to ~$150k of annual normalized costs by 2022). The collaboration has shown the ability to use a variety of resources over prior years. In 2019, primarily FermiGrid resources were used. Additionally, in 2019, resources at Argonne were used (including Theta); these were run outside of HEPCloud and were approximately twice the reported FermiGrid usage.

MicroBooNE:

Data-taking operations of the MicroBooNE experiment is assumed to have ended. If there is a special reason (such as a tantalizing finding in the data) that might trigger a new request to the Lab by the collaboration to take some additional data.

The experiment had a significant increase in tape usage beyond the estimate in 2019. For year-to-date figures, it is slightly over the estimate. Approximately 70% of the estimated costs of supporting MicroBooNE is in the tape usage.

NOvA:

NOvA collaboration has an organizational structure for offline computing but has no liaison with the computing sector.

NOvA's previous cpu requests and actual usage have been pretty close. The requests for future years look reasonable. Their tape and disk requests also are commensurate with size and scope of the experiment, and are justifiable.

The collaboration has had issues with OSG usage. Had dedicated personnel in past years but wasn't able to keep this up. They would like better monitoring tools and resources to scale up on OSG again. The collaboration also has not been able to exploit HEPCloud because of lack of resources.

The collaboration reported that they had insufficient hours this year on NERSC for their compute-intensive analysis application of performing the Feldman-Cousins (and other) calculations of confidence intervals and significance. Would be interested in porting this application to other HPC facilities. NOvA will also benefit from GPU resources for their ML applications in neutrino event reconstruction and classification.

SBN:

SBN encompasses ICARUS and SBND, with ICARUS planned to begin data taking this year and SBND the following year. This seemingly drives the tape usage from < 1PB to nearly 50PB required at the end of 2022. It appears that about half the request is derived data and the experiment has set an internal goal of deleting derived data from tape after 2 years.

Muon g-2:

Run 3 of the experiment has just completed. Run 4 begins in 2021 and will double the existing dataset. Core-hour usage of SCD resources will increase from 15% to 23% of currently available resources by 2022. Persistent (dedicated) dCache is predicted to increase from 4% (2%) to 7% (2%) of total user requests. Tape usage is expected to remain at about 15% of the total user request through 2020 (increase from 11PB currently to 30PB). The experiment reports methods of data organization that allows deletion of old sets as new releases become available.

CMS:

CMS is ~50% of FNAL computing resources and works closely with the HEPCloud team to add opportunistic cycles. Expecting a 30% increase in core-hours of compute, 30% increase in tape storage, and 15% increase in disk storage demands in 2021 over 2020.

DES:

DES uses a negligible amount of SCD compute resources (roughly 2% of annual core-hours), a moderate amount of tape storage (5% of current total used), and a significant amount of persistent dCache (25% of current total used). The predicted increase is not significant with 0-5% increases across all resources in the next two years.

LSST:

LSST resource usage of SCD resources has been small and so far counted in with DES. Compute usage and storage requirements are not predicted to exceed the 1% level through 2022.

Other Experiments:

Minerva completed taking data by the end of February 2019 but will continue analyzing data through 2022. It is the largest user of SCD resources out of the 'other' experiments. Minerva currently represents about 7% of core-hours available in SCD and this is expected to remain the same through 2022. Persistent (dedicated) dCache usage is currently around 11% (6%) of current total facility usage. The absolute amount used is not expected to increase over the next two years. Tape usage is expected to not exceed 2% of total facility usage over the next two years.

SeaQuest/SpinQuest uses mostly OSG opportunistic core-hours and uses <1% of the available core-hours in SCD. This is expected to remain the same in the next two years. dCache usage is expected to stay at or below 3% of current facility usage in the next two years. Tape usage is expected to remain <1% of current facility usage through 2022.

MINOS currently uses 2% of total facility capacity in core-hours and this is expected to fall in the next two years; dCache usage is expected to remain at the current level of ~2% of current facility usage for the next two years and tape usage is expected to stay at or below the current level of ~1% of current facility usage.

MiniBooNE does not use core-hours or dCache in SCD, and is predicted to store on tape <1% of current facility usage.

Comments:

General Comments:

- Despite the change in this committee's name, the process this year was essentially the same as in previous years including the charge given to the committee. It was also felt that the process does not afford real "scrutiny" of the resource requests, since there is not enough time or sufficient interactions with the experiments to judge if the requests are justified.

- The CERN CRSG process is very rigorous and involved. Since there is the stated desire to make the FCRSG activities similar to the LHC's CRSG, adapting several aspects of the FCRSG process should be explored. The level of scrutiny could be adapted based on the level of resource requests. For example, if an experiment's resource request requires more than $500k of annual computing costs, then the experiment could be paired with a member of the review committee who would act as an "auditor" to do a thorough review of the resource estimates before the FCRSG review.

- We commend SCD for adopting and using Rucio as a common data management framework. That being said, last year there were discussions of common development projects such as HEPCloud and ART. This year there was little mention of these tools. It was unclear if this reflected a real change in priorities or some other effect.
- The recent SCD staff reduction is concerning. This might lead to an unplanned scope decrease as expertise is lost randomly and not replaced. We are concerned that attrition does not protect critical capabilities during staff reductions, and it can affect some projects and core competencies negatively.

Computing Services

- Last year's CPU requests were not very good estimates for some of the experiments. However, since some experiments overestimated and others underestimated, the usage was manageable overall. The committee does not see strong incentives for the experiments to try to fit within their estimates, and so they are not naturally motivated to improve their efficiency.
- There is some fraction of idle on-site CPUs and there are several contributions to this. For example, in previous years, focus was on jobs idling while waiting on tape staging, while this year the main issue seems to be jobs over-requesting memory (leading to idle cores with insufficient free memory to start new jobs). The SCD can take a holistic approach: additional throughput can be gained by avoiding on-demand staging, by working to decrease memory footprints, and (potentially) by encouraging experiments to improve the accuracy of their memory requests. SCD could use its monitoring tools to discover the source of inefficiency and allocate effort appropriately to tackle problems.
- SCD should be a resource for the experiments and stay ahead of the curve in testing architectures, but experiments must be the ones to optimize their algorithms.
- Fermilab is working to leverage large-scale external computational resources (e.g., DOE Leadership Computing Facilities) while maintaining data repositories on-campus. While SCD identified a future strategy for maintaining data repositories and data management services on-site, the storage services presentation identified no R&D efforts toward this goal.
- R&D efforts to date have focused on porting applications to run efficiently on new computing architectures (including through projects receiving CCE funding) and on extending the Fermilab infrastructure to overlay other resources (HEPCloud). For HEPCloud, no plan was presented to allow access to resources beyond NERSC, while the division has a goal for 2020 to utilize Theta at Argonne.

- Beyond porting of algorithms to accelerator platforms, there are significant infrastructure challenges in utilizing HPC centers and data management. Particularly, there's an architectural change when the data and computational resources are at different sites. HepCloud could be more effectively used as a gateway to these resources.

Storage Services

- It was noted that most experiments do not have clear policies for data lifetimes. Some are storing very large data samples but cannot articulate the value of keeping them in the tape library or on disk storage, at significant costs. For example, a significant portion of the data on NAS is greater than 4 years old (at odds with the mission for this resource). The lack of a data lifetime policy is most significant for tape storage -- for several years it has been noted that the costs of migration between tape media generations incurs significant and ongoing yearly cost.

- The committee heard that tape costs would go up by about a factor of two to fulfill all experiment requests by 2022, and it was stated, "tape requests are potentially on an unaffordable trajectory." Cost for tape storage makes up a significant fraction of the overall resource needs, for example it is ~70% of the requested computing cost for MicroBooNE. While tape is driving the cost for storage, several of the large tape requests in 2019 were incorrect by multiple petabytes. MicroBooNE and DES used almost twice their requests while a few others used much less than requested. Such large errors in estimates render overall planning ineffective.

- Experiments expressed little concern for being over-budget. In 2019, the tape overuse by several experiments did not lead to an overall crisis only because the SBN/ICARUS experiment did not yet take data. Also, it was not clear to the committee how much tape usage are Raw (must-keep) vs Reco (derived). Further follow-up to improve the computing resource models of experiments may be justified.

- As discussed last year, it's unclear whether the current use of dCache as a hierarchical storage manager (allowing on-demand staging) is viable for some experiments. It's noted that larger experiments and sites have abandoned this model for several years now.

- It also appears that the dCache resource requests are often vaguely justified. For example, the dCache dedicated space requested for DUNE grows by 1.5PB over the next 2 years (from 25% of the total in 2020 to 60% in 2022) and no explanation was given.

Recommendations:

1. The committee would like its mandate to be more specific and clearer. If the Charge for the group would be to scrutinize experiments' resource requests only and not the SCD vision/plans, then that should be clearly reflected in the Charge.
2. The committee suggests that the level of scrutiny be based on the experiment size with a threshold on cost.
3. The committee asks that common computing capacity units are defined and used for easier comparisons, i.e. core-hours used, bytes stored, etc., instead of cost on the cloud. The units should also be presented as a percent of total facility usage and percent of total facility availability for each experiment.
4. The committee recommends that the SCD require experiments to develop data lifecycle policies. SCD should institute a default (finite) data lifetime for all derived data written to tape. Experiments should be able to request changes from the default - but given the significant recurring yearly costs of tape, extraordinary requests should be well justified.
5. Computing architectures are changing very fast. The SCD should allocate resources for testing new architectures and algorithms. This may be in the form of on-site testbeds or arranging partnerships with other DOE labs or CERN openlab.
6. The committee recommends that the SCD work with the experiments and strategic partners on demonstrating the ability to design and test new architectures at scale.
7. The HEPCloud represents significant long-term development expenditure. The committee recommends exploring extension of its use for connecting to HPC sites as a common gateway.
8. The SCD should facilitate the development of AI/ML implementations through hardware resources and data science expertise. SCD should define its objectives in AI/ML and develop a strategic plan that will enable it to become a leader in AI/ML for the field, and to support application of AI/ML across the Lab and the community.

Response to Charge:
(As presented at the Closeout)

Evaluate Challenges

1. DOE has reduced operations funds
    a. SCD staff reduced by 10 (by attrition) during the past year.
    b. Requests for SCD resources are growing nonetheless
    c. SCD must prepare for the upcoming generation of experiments while also adapting to the changing computing landscape

Comment: We are concerned that attrition does not protect critical capabilities during staff reductions; it can affect some projects and core competencies negatively.

2. SCD was already operating in "lean mode" before the most recent cuts
    a. Funding has not supported basic refresh of CPU resources. Some new hardware purchased in FY19.

Comment:

It appears the onsite CPU resources are largely sufficient to meet the requests over the next two years despite insufficient funding for basic refresh. However, this only satisfies the 'baseline' needs and doesn't allow experiments to chase opportunities or to shift focus to the future (see next comment).

Comment: We think that SCD has a general idea what the challenges of the future are in terms of needs for future experiments and changing technologies, but it is less clear how to succeed in these, e.g. sufficient effort and talents to participate in testbeds, and evaluate new technologies.

Evaluate Requests

1. Are the experiments' resource requests believable?

Yes, overall. Based on past performance, the resource requests for CPU are plausible. In absolute terms, the errors for the CPU estimates in 2019 were within what could be absorbed by available resources (dedicated on- and off-site, allocations, and opportunistic). This year, there are far fewer large jumps than in prior years. However, some of the storage requests were off by large factors, and given the large tape storage cost this has a large impact on planning.

We will provide more detailed experiment-specific comments in the final report.

2. Are the experiments' resource requests justified?

The requests make sense. Although, the information and the process precluded us from determining whether resource requests are fully justified.

If a detailed scrutiny is needed for the larger experiments/projects, more exposure to the assumed models and physics goals will be needed.

Comment: Fully burdened costs do not allow comparison of resource requests across years and they also hide implementation details and facility efficiency. The fully burdened costs would be useful to the experiments so they know the relative costs of technical choices. In terms of scrutinizing resources requests, standard metrics may be more useful to use.

3. Is the SCD response appropriate? Especially if requests > resources?

Yes.

SCD was fortunate this year to have been able to respond even to overages. It is not clear that there is currently a mechanism to limit users/projects resources.

We note few experiments had presented a lifetime model for data on tape. Given that archival capacity drives the increase in cost for the next few years, this needs urgent work. SCD should consider setting a default, finite, lifetime for data stored and require experiments to specify (and justify) changes from the default.

4. Is there room for more optimization?

Yes, but it is not clear what the trade-offs are on the experiments' side and the SCD side.

There might be optimization on what data gets stored and what can be reproduced at a lower cost has to be looked into. SBN, for example, plans to store significant derived data on tape despite the fact the CPU cost to derive the data appears lower than the tape cost.

Currently CPUs need to be undersubscribed to provide enough memory for some fraction of the applications. This is a measurable efficiency loss that may be recovered by multiple potential means (examples: more accurate memory requests, multicore jobs, reduce memory usage).

5. Can the committee provide specific feedback to experiments/projects?

Yes. We will give specific feedback on experiments in the committee's full report. Here we want to point out a couple of examples that are of high importance. HPC and the challenges of data management to use external resources will be critically important for the SCD strategic goals. We encourage international collaboration and strategic planning in these areas.