FIFE Notes - October, 2015 News

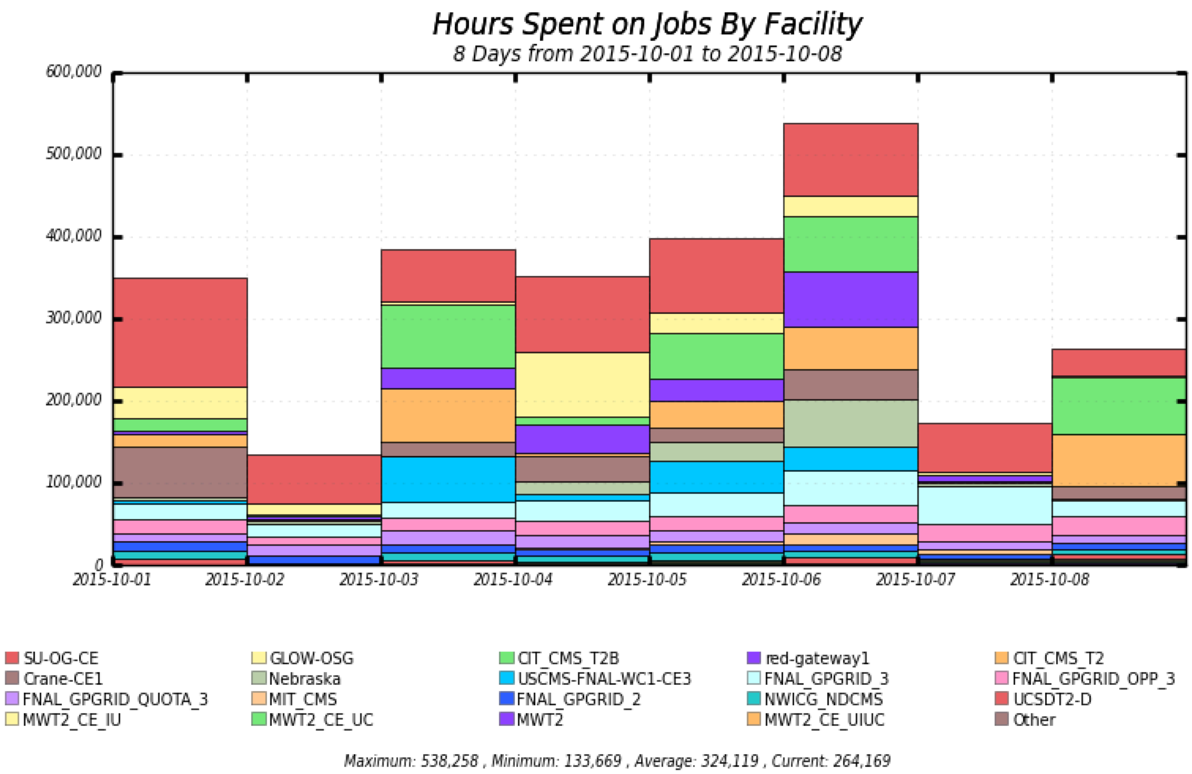
for

Distributed Computing at Fermilab

Mu2e Simulation Campaign:

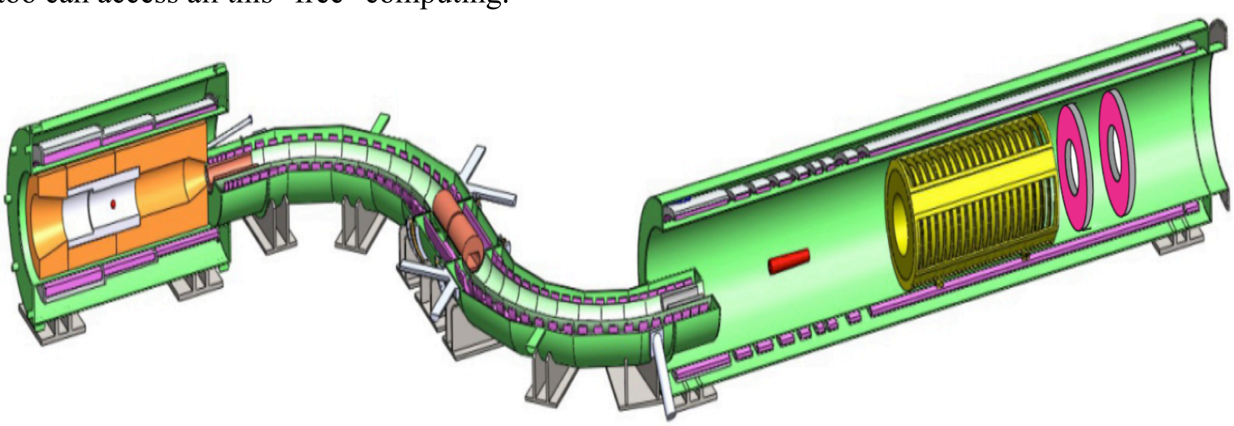
Utilizing the Open Science Grid to increase sensitivity to new physics

When designing an experiment, one of the greatest challenges is having confidence that you have optimized the sensitivity of your measurement, squeezing every last bit of precision out of your resources. The Mu2e experiment is preparing to go through DoE Critical Decision 3c review next summer and they face this challenge when designing their apparatus including the solenoid system, shielding, and detector components. In order to model the various beam transport options, detector designs, and estimate background rates, the experiment estimated they would need to run millions and millions CPU hours of simulation in the months leading up to their review. This was well beyond the Mu2e allocation of computing resources at Fermilab, and so Mu2e turned to the Open Science Grid for opportunistic CPU hours.



The daily CPU hours utilized by the Mu2e experiment on the Open Science Grid during Oct 1 - 8, 2015. The peak utilization is greater than 500,000 CPU hours in a single day with the average CPU usage for the work being equivalent to 13.5K cores of computing which is more *than the entirety of Fermilab GPGrid resources.*

The simulation and calculations needed by Mu2e fit perfectly within the model of OSG operations. The calculations had relatively little input (small text files to configure the simulation parameters), heavy CPU utilization (modeling the transport of millions of particles through the solenoid), and small output files consisting of art-event root files of ~100 MB per job. With the help of the FIFE group, Mu2e was able to distribute their simulation code to worker nodes with CVMFS, and then opportunistically run jobs on 15 different OSG sites across the country. Mu2e was able to utilize more than 20 million CPU hours in only 5 months, averaging about 8000 concurrent jobs with peaks usage as high as 20,000 simultaneous jobs, all without any direct capital expenditures from Mu2e. Mu2e has already met their base goal and this successful campaign continues as they work to achieve their stretch goal of simulating 24 times their planned operational live time for three critical backgrounds. The results of these calculations have drastically increased the understanding of the experiment design well beyond anything what might have achieved with Fermilab CPU resources alone. If your experiment, proposal, or idea needs more computing resources from the OSG, please contact the FIFE group to learn how you too can access all this "free" computing.



Layout of the Mu2e Apparatus from http://mu2e.fnal.gov/public/gen/

- Mike Kirby

Introducing OPOS

Photo courtesy of Luis Contreras (Fermilab in 2015). Initial OPOS team consisted of (from left to right), Qiulan Huang from China, Jeny Terheran from Colombia, Anna Mazzacane from Italy, Vito Di Benedetto from Italy, Paola Buitrago from Colombia. Since this photo was taken, Jeny has moved to the security team and Andrés Felipe Alba Hernández from Colombia has joined the team. Felipe is pictured below.

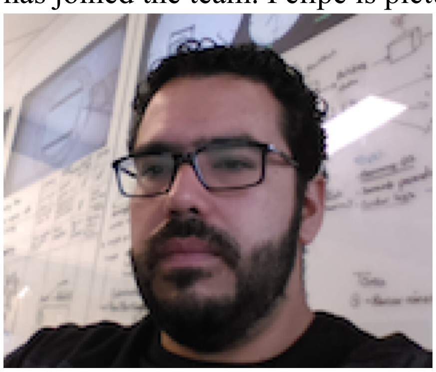


Photo courtesy Andrés Felipe Alba Hernández

Offline Production Operations Service (OPOS) helps transform the raw data collected by detectors into a format that scientists can interpret so they can give it a physical meaning. OPOS does the behind-the-scenes work needed to perform the analysis of the data, an important step to contribute to scientific publications. The scientific community shares knowledge and provides the base for future research by publishing their research.

The OPOS team contains skilled computing scientists and physicists from different regions of the world. Their combined expertise assists Fermilab experiments in running their large-scale production data. The OPOS team was created following the successful example of the CMS Operators team. Instead of one experiment, OPOS has the challenging task of multiple experiments with different lifecycles, needs, tools, deadlines, etc. Small collaborations are often short of staff, so the OPOS expertise expands the human resources of the experimental offline team.

The OPOS group constantly assists experiments and contributes to improve their production efficiency. This "new born" team is ready to make a thoughtful and impactful contribution at the lab. Stay tuned!

- Anna Mazzacane

News from the OSG

FIFE CVMFS Repositories Now Distributed Worldwide

There are currently 14 CVMFS opensciencegrid.org repositories hosted at Fermilab that are available to the Open Science Grid , to all sites that have an OSG installation of the cvmfs client. There are now also five repositories -- fermilab, lsst, mu2e, nova, and uboone -- that are available to any site worldwide that has installed a standard cvmfs client from CERN or EGI in the last year.

CVMFS clients read data from Stratum 1 servers, and there are five of those worldwide in the default client configuration for the opensciencegrid.org domain (see map). Two of them -- at Fermilab and BNL -- are also used within the OSG, so all opensciencegrid.org repositories are on those. Worldwide clients could read long-distance from those, but in addition to increased latencies, due to a limitation of the cvmfs automated ordering randomly choosing 3 servers to calculate an order, sometimes those mounts will fail.

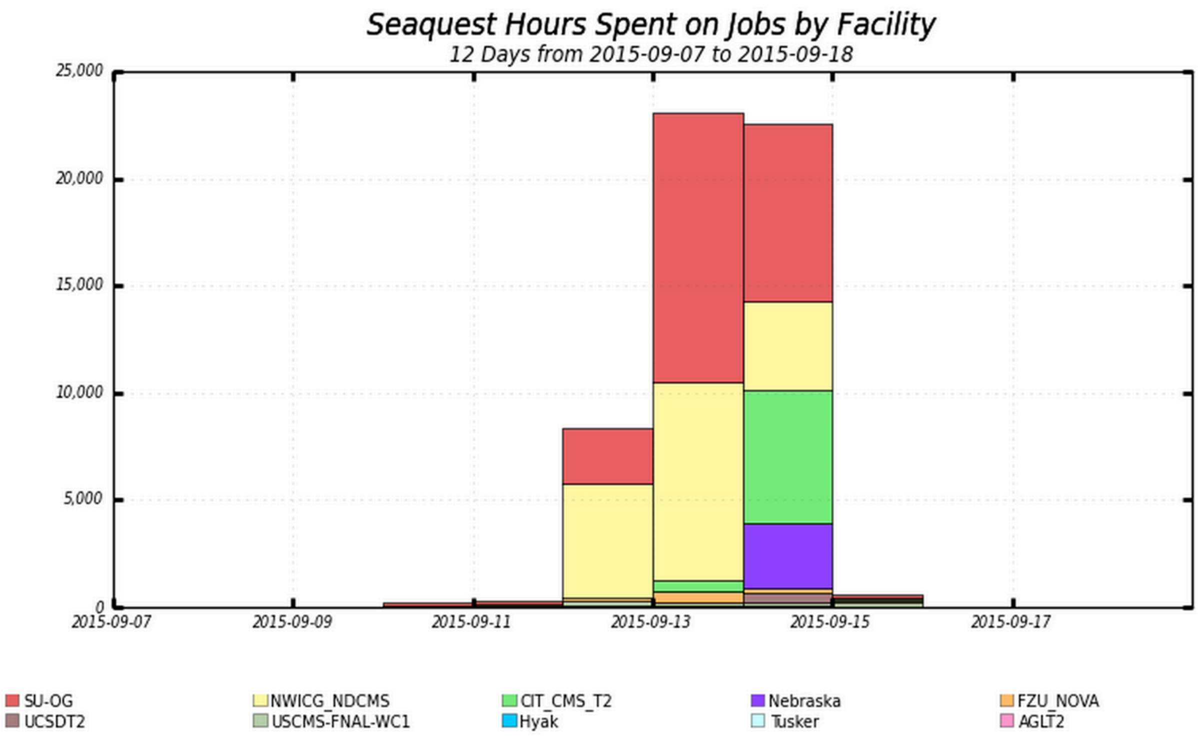
EGI has a documented process https://wiki.egi.eu/wiki/PROC20 for importing opensciencegrid.org repositories to the other 3 Stratum 1s. EGI first requires that a VO be accepted by at least two EGI sites. The Fermilab VO has been accepted by EGI, and most of the FIFE VOs are considered to be "sub-VOs" of the Fermilab VO, so if any of the others want to make their repositories available worldwide, they can be fast-tracked through the process when they request it in a GGUS ticket and say they are asking on behalf of the Fermilab VO.

- Dave Dykstra

Helpful hints

A SeaQuest tale: the Quest for Open Science Grid

SeaQuest was in a difficult situation this summer: a looming deadline for two fall conferences, a big crunch for data processing in preparation for them, the need for final tweaks to the tracker code with pressure on the few experts that could make it all happen. It became apparent that FermiGrid would not have enough opportunistic cycles to finish this computation. SeaQuest was running out of time and resources to deliver on key physics results. FIFE helped SeaQuest use resources available on the OSG.



SeaQuest computing hours on OSG – Graph from http://gratiaweb.grid.iu.edu/gratia

With FIFE support, SeaQuest was successful in running on OSG. Overall, the experiment ran 1,100,000 CPU hours for 185,000 jobs, transferring 7.5 TB of input data and 9.5 TB of output. About 10% of the computing was executed at 10 OSG sites, running for the last 3 days of the campaign at the level of 500-1,000 CPU slots continuously, with an overall 90% efficiency.

This computational campaign supported SeaQuest in achieving its physics goals for the fall conferences. The SeaQuest experiment continues a series of Drell-Yan measurements to explore the antiquark content of the nucleon. Measurements from high-energy scattering experiments are consistent with quantum chromodynamics (QCD) theory predictions where a sea of virtual gluons arises in the nucleon; these gluons radiate other gluons or pairs of quarks and antiquarks. Gluon splitting, for example into an u anti-u quark pair or a d anti-d quark pair, is a flavor symmetric process. As such, statistically the amount of anti-u and anti-d quarks is the same. However measurements have shown that this amount differs by up to 50% indicating that processes other than gluon splitting contribute to the nucleon sea. Extending the existing measurement of the anti-d / anti-u ratio to kinematic regions where the ratio has not been constrained is a key measurement of SeaQuest. This will help identify theories effective at describing the nucleon sea and exploring the origin of the nucleon sea.

The contribution of the Scientific Computing Division was crucial to enable SeaQuest to meet its deadlines. Particularly effective was the help of FIFE support in porting the computation to OSG by troubleshooting jobs, identifying areas in the code that needed modifications to be compatible with the OSG environment, and helping set up the software distribution system.Not only did SeaQuest obtain results in time for the conferences, it also learned how to run on OSG, a practice that will be useful for the computing production campaigns to come.

- Gabriele Garzoglio

Service Updates - Monitoring: Why are my jobs not running?

In the course of using a batch system, it is inevitable that users ask this question. Some solutions include opening a SNOW or GOC ticket, sending an email, pinging someone on IM, or dispatching a carrier pigeon to ask the admins to look into it. This quickly becomes untenable in a system with hundreds of users and tens of thousands of simultaneous jobs running, such as Fifebatch. Enter [Fifemon](https://fifemon.fnal.gov:3000/), the near-real-time monitoring system for Fifebatch, where users can see at a glance the status of their jobs, their experiment's jobs, and the status of the Fifebatch system as a whole, including local and offsite grids. From this assemblage of data, users can quickly answer the question of why their job isn't running. They can ascertain what action may help get their jobs running sooner. At least that's the vision, one we are quickly developing into reality.

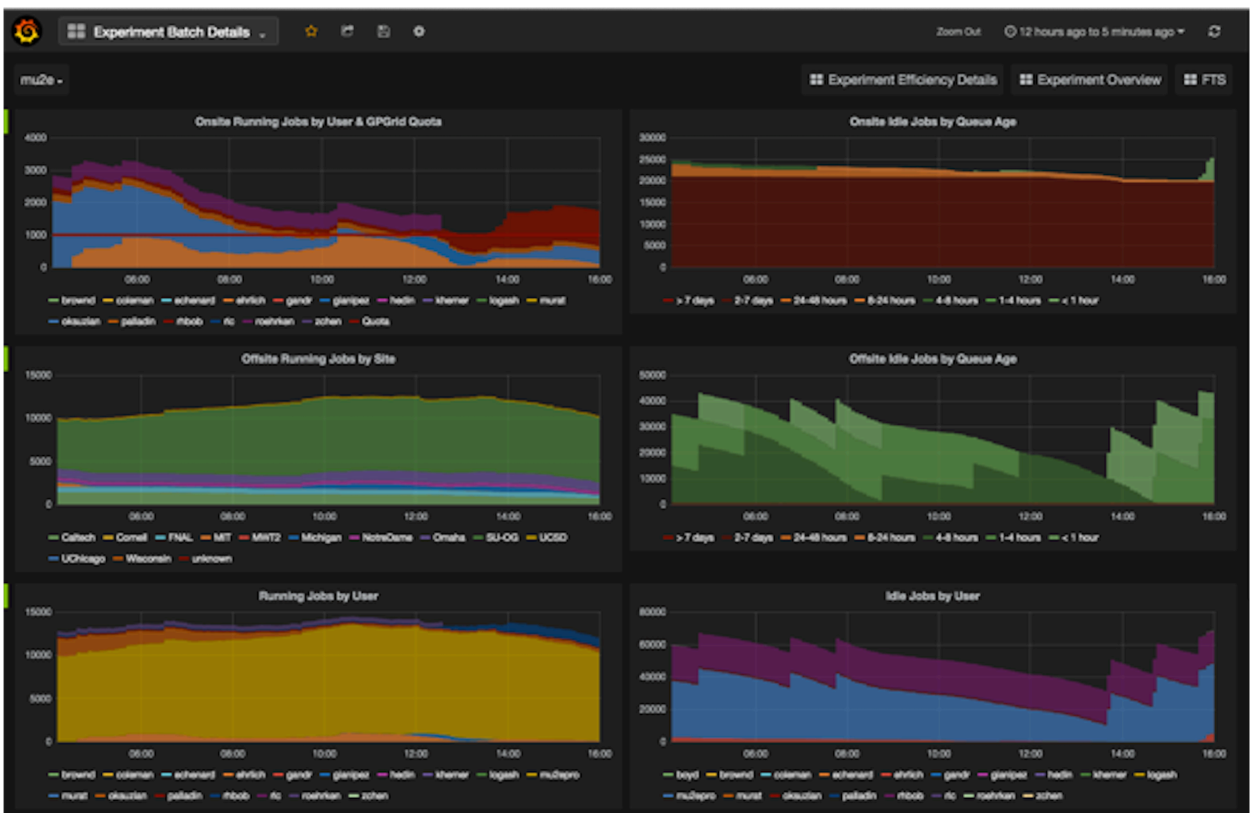


Photo from Fifemon on the Mu2e on 9/30/15: https://fifemon.fnal.gov:3000/dashboard/db/experiment-batch-details?var-experiment=mu2e

As this newsletter goes to press, the Fife team is in the process of releasing the next evolution of Fifemon, which is running on the open-source Graphite and Grafana packages. Adopting this framework allows us to focus effort onto what data is collected and how it's presented. We can rapidly iterate on this. We are tracking over fifty thousand metrics related to the Fifebatch system, a number that increases every week, and we are constantly exploring new and better ways to present this data in a meaningful way. We plan to eventually open the system up to you, to develop your own views, to allow you to focus on the data that you care about! In the meantime, please explore the dashboards and graphs, and let us know what else you want to see (please don't use carrier pigeon, SNOW tickets are less messy.)

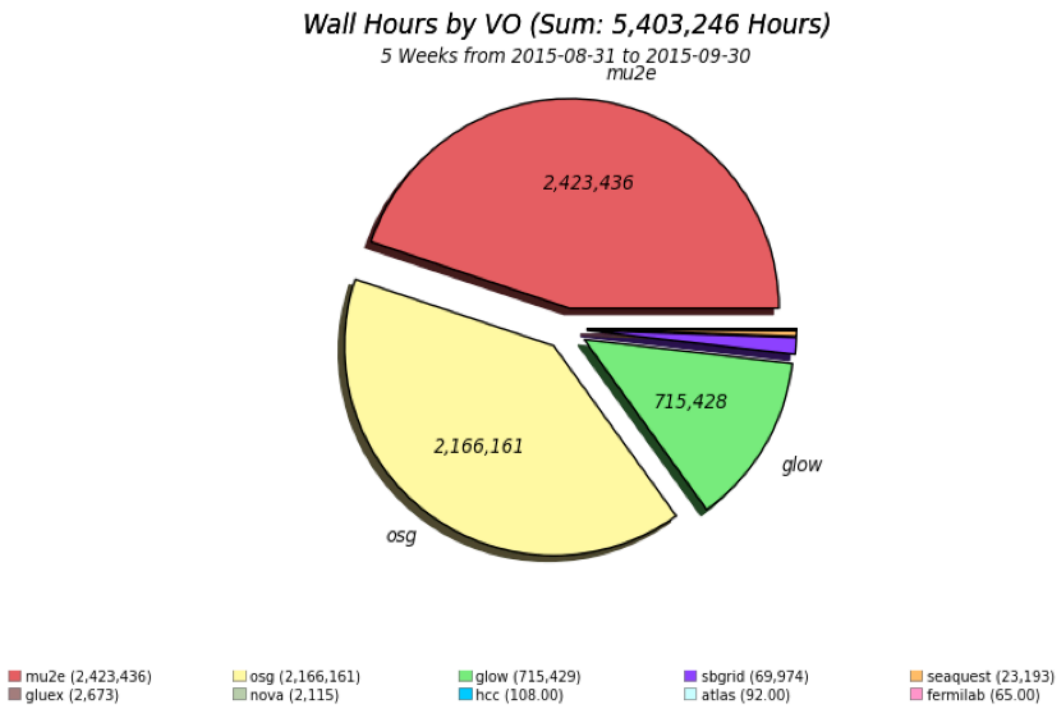
- Kevin Retzke

Most opportunistic cycles for a site

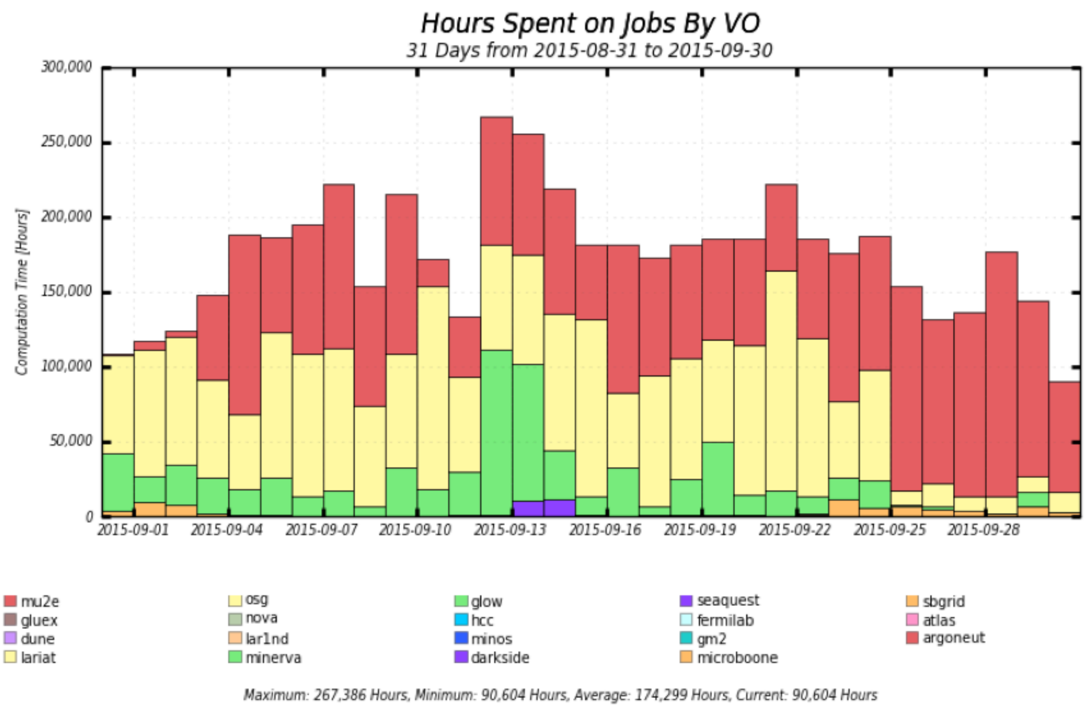
In the month of September, sites on the Open Science Grid (OSG) provided 12 million computational hours to opportunistic users—users who do not own any of the resources on the site. The single largest source of these resources was the Syracuse University OrangeGrid (SU-OG), which provided over 5.4 million hours. SU-OG, developed in 2012, harvests idle computing power in computers across the Syracuse University campus for research computing and has now grown to include over 12,000 CPU cores. In November 2014 SU-OG was added as an OSG site, allowing capacity unused by Syracuse users to go to users and virtual organizations (VOs) of the OSG. In the month of September, the VO harvesting the most cycles from SU-OG was Mu2e, which obtained 2.4 million computational hours.

More about SU-OG: http://researchcomputing.syr.edu/resources/orange-grid/

OSG Accounting: http://gratiaweb.grid.iu.edu/



Graph from: http://gratiaweb.grid.iu.edu/gratia/



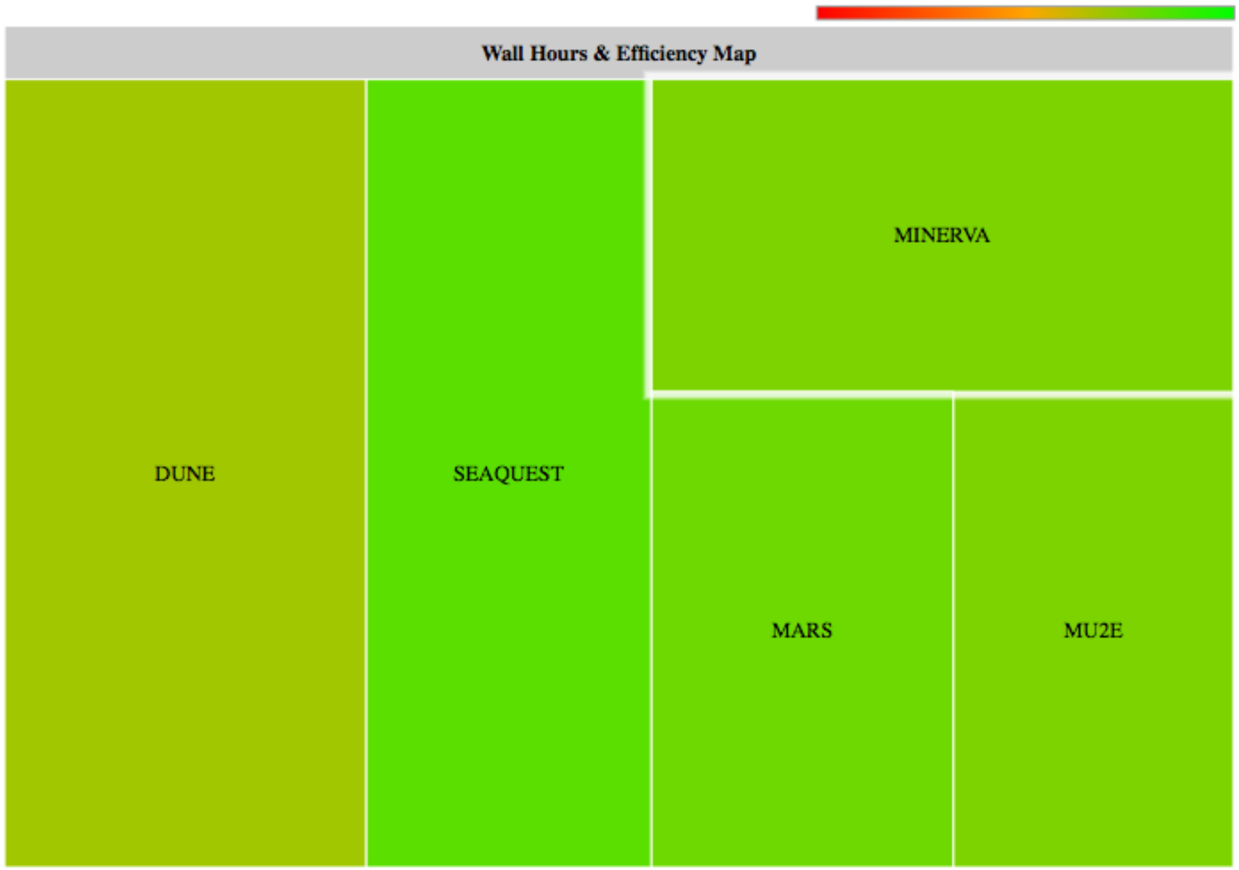
Graph from: *http://gratiaweb.grid.iu.edu/gratia/*

- Bo Jayatilaka

Most efficient experiment

Most efficient experiment on FermiGrid that used more than 500,000 hours since August 1st -- SeaQuest : 82.22%

(The brighter the green in the following chart, the more efficient.)



Report from http://fermicloud035.fnal.gov:8100/gratia/

- Tanya Levshina

Most efficient big non-production user

Most efficient big, but not Production, users on FermiGrid who used more than 100,000 hours since August 1st was Laura Fields from DUNE with 98% efficiency.

Experiment	Name	Hours	Efficiency
MINERVA	Jeremy Wolcott	172,705	0.97
DUNE	Laura Fields	139,312	0.98

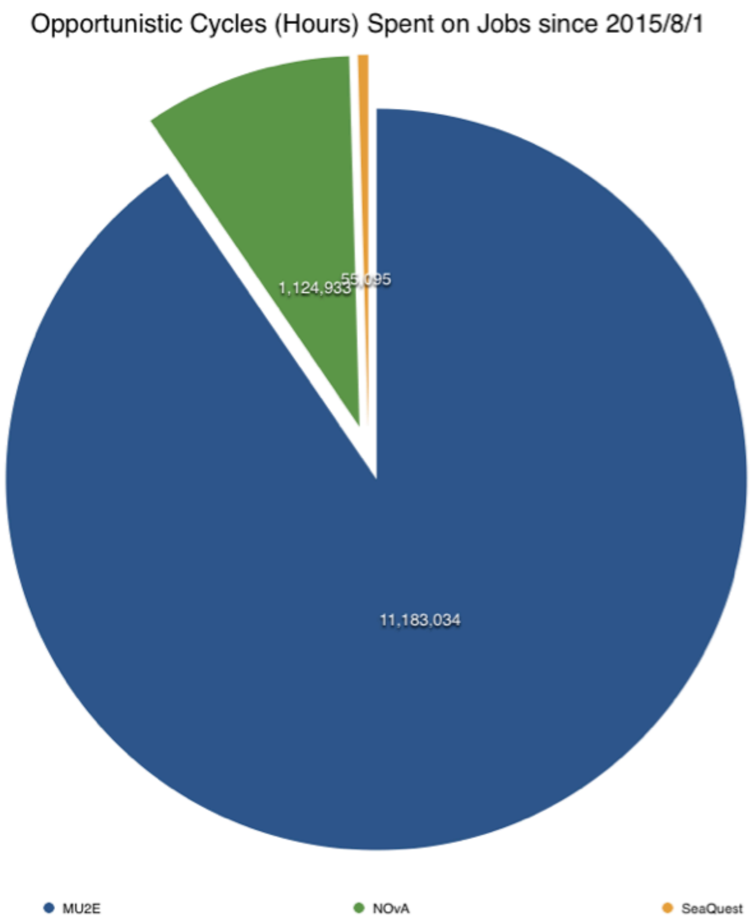
- Tanya Levshina

Experiment with the most opportunistic hours

The experiment with the most opportunistic hours on OSG between August 1 and September 30 was Mu2e with 11,183,034 hours.

Top three:

- 11,183,034 - Mu2e
- 1,124,933 - NOvA
- 5,5095 - SeaQuest



- Tanya Levshiva

Brought to you by

This newsletter is brought to you by:

- Dave Dykstra
- Gabriele Garzoglio
- Bo Jayatilaka
- Mike Kirby
- Katherine Lato
- Tanya Levshina
- Anna Mazzacane
- Kevin Retzke

Feedback

To provide feedback on any of these articles, or the Fife notes in general, please email fife-group@fnal.gov