NERSC Usage with GlueX Raw Data in AY2018

David Lawrence

October 9, 2018

Introduction

The first use of NERSC to process experimental raw data from GlueX was done in 2018. This marks the first time an offsite computing facility was used to process a large set of experimentally derived data from JLab. This document describes some details on what was done, and how we might expect to use the facility over the next year.

In the Spring of 2018 GlueX produced a raw data volume of approximately 1.5PB. In the previous year 1PB was produced and in the Fall of 2018 another 1.5PB is expected. The data is processed in multiple passes with the first ones covering only a small subset (2%) to test calibrations and software changes. The output of a single pass produces roughly 1/3 of the input data volume. Thus, significant data volume needs to be transported both to and from NERSC. The first campaign done at scale was "monitoring launch ver17" of the Spring 2018 data set (i.e. RunPeriod-2018-01) consisting of $\approx 100TB$ of raw data. Approximately 30TB of output was produced in this campaign and transferred back to JLab. Subsequent smaller scale tests were done to exercise individual components of the system and are also reported here.

1 Reconstruction Software Preparation

The GlueX reconstruction software is continually evolving with frequent changes leading to new version tags. This motivates us to use a system that can easily distribute newly compiled versions which can then be run on all production platforms, regardless of the host OS. This is done using a Docker container to provide cross-platform uniformity and CVMFS for binary distributions. NERSC supports the use of containers via it's SHIFTER system[1] which can read easily read from Docker. The standard Docker container used for both OSG and NERSC running by GlueX is:

---

[1] https://docs.nersc.gov/development/shifter/how-to-use

docker:markito3/gluex_docker_devel

The CVMFS volume used is /cvmfs/oasis.opensciencegrid.org/gluex. This mirrors sections of the */group/halld* disk on the JLab CUE. Thus, new builds at JLab are automatically distributed to offsite facilities via CVMSFS. The Calibration Contastants DataBase (CCDB) is also distributed via CVMFS in the form of a SQLite file. Similarly for "resources" like the magnetic field maps.

2 NERSC Performance Testing

2.1 Monitoring Launch RunPeriod-2018-01 ver17

The first production test of swif2 and NERSC was done over 12 days in August 2018. This was recorded as the RunPeriod-2018-01 ver17 monitoring launch[2]. During this campaign, the first 10 files for each run with "@is_2018production and @status_approved" were processed for a total 5,495 files. Each file was 20GB and submitted to swif2 as a separate job using the *offmon_2018-01_ver17* workflow in the gxproj4 account on the JLab CUE. Scripts used to form and run the jobs were placed in the same github location as those used for processing data on the JLab SciComp farm using swif.

Jobs were submitted with a limit set in swif2 to only allow 1000 jobs maximum to be in-flight at any given time. Figure 1 shows the number of jobs queued (blue) and active (red) at any point in time during the campaign. Some issues with swif2 during the first 1/3 of the campaign prevented the queue from reaching 1000 jobs during that time. The number of simultaneous jobs peaked at round 300, but the average was 62. The overall latency between submitting the job to the slurm system on Cori I to the job completing can be seen in figure 2 using data mined from slurm. The NERSC webpage showing the queue backlog during this same time period as a function of date can be seen in figure 3. According to NERSC, the backlog during this time was approximately 7-12 days.

2.2 ESnet Bandwidth Testing

Figure 4 shows a screen capture of the ESnet bandwidth usage during the RunPeriod-2018-01 monitoring launch ver17 done via swif2. The spikes above 2Gbps are likely due to transfers done by the JLab lattice group also using NERSC at the time. Transfers initiated by swif2 appeared to be limited to 2Gbps aggregate. All files were 20GB and swif2 transferred each one via a separate Globus transaction (5495 total). A total of 110TB was transferred from JLab to NERSC during this campaign and files totaling roughly 1/3 of that

---

[2]Details can be found on the launch analysis website here:
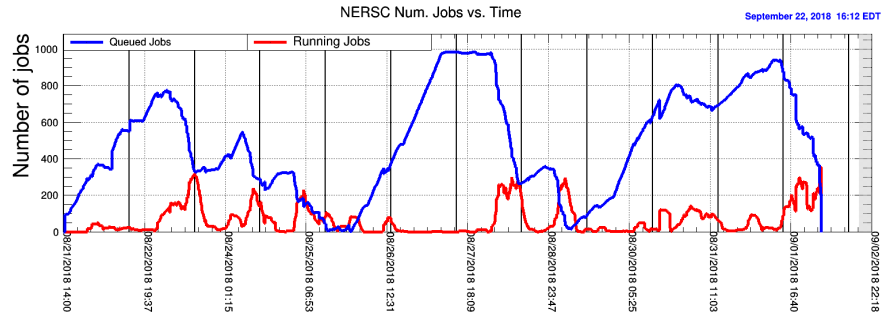*https://halldweb.jlab.org/data_monitoring/launch_analysis/index.html*

Figure 1: Instantaneous count of number of jobs in queue and number of running jobs on Cori I(haswell) during the RunPeriod-2018-01 ver17 monitoring launch.
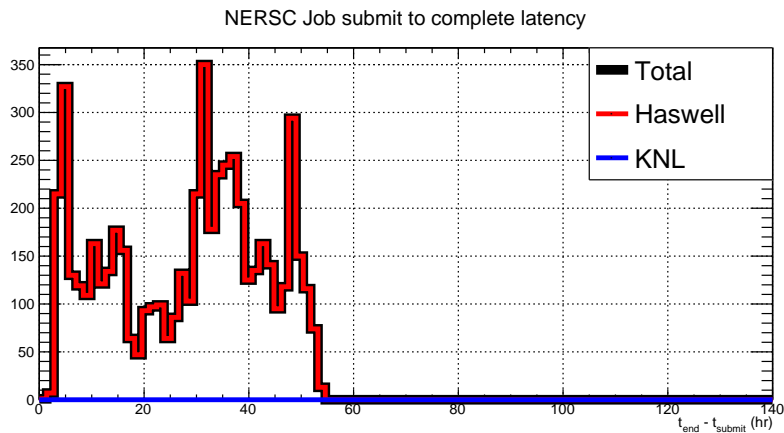

Figure 2: Latency between submission time and job completion time for Run2018-01 ver17 monitoring launch on Cori I (haswell). Times also include the time for the job to run which was 3.6hr on average.

Figure 3: Backlog of all jobs on Cori I (haswell) during the time the Run2018-01 ver17 monitoring launch was performed.
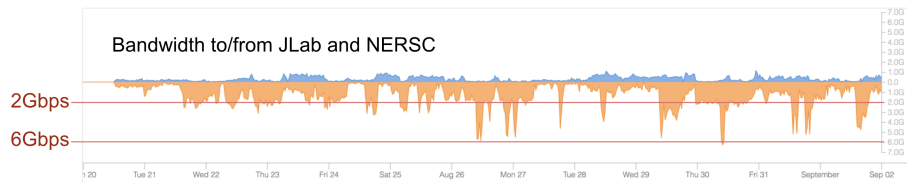

Figure 4: Screen capture of bandwidth usage between JLab and NERSC during the time of the RunPeriod-2018-01 moniroting launch ver17. Some activity was due to other groups (HPC) which generally resulted in the spikes over 2Gbps.

were transferred back to JLab. It is worth noting that during this time there were routinely dozens of pending transfer requests so the 2Gbps limit was not due to using a single stream.

Figure 5 shows a test where files from 3 runs totaling 14TB were transferred to NERSC via globus. In this test, 3 transactions were submitted (1 for each run). Each consisted of 232 files, 20GB each in size. The average bandwidth achieved for this test was about 6.5Gbps with peaks over 10Gbps (the hardware bandwidth limit).

Figure 6 shows a series of 3 tests where multiple numbers of transactions were submitted to transfer raw data files from JLab to NERSC. These tests indicate that at least a few streams are needed to fill the bandwidth. However, the 4 stream test did not seem to achieve the same average bandwidth of the earlier 3 stream test shown in figure 5. The reason for this is not known.

Further testing with swif2 using a new feature to bundle transfers will be attempted this Fall. It is expected to reproduce the higher throughput results shown in these tests. Future upgrades to the JLab ESnet connection which are

Figure 5: Screen capture of bandwidth usage between JLab and NERSC during a bandwidth test. During this test, 698 files, each 20GB in size was transferred from JLab to NERSC using 3 transactions of 4.6TB each (total of 14TB).
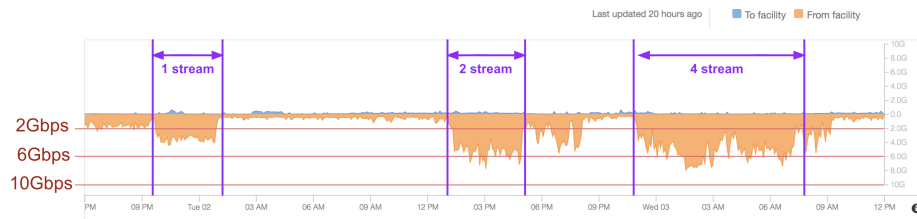

Figure 6: Screen capture of bandwidth usage between JLab and NERSC during the time of the RunPeriod-2018-01 monitoring launch ver17. Some activity was due to other groups (HPC) which generally resulted in the spikes over 2Gbps.

expected to be phased in over the next 2 years should provide a greater ability to ship data offset. This will be needed as GlueX begins high intensity running which will produce larger data sets ($\approx 5PB/yr$).

2.3 NERSC Latency Testing

Over the time period covering July to October 2018 several exercises were done to test the latency of jobs and develop a workflow for using Cori I and II at NERSC. This included the RunPeriod-2018-01 ver17 monitoring launch done through the swif2 system. Figure 7 shows the latency between a job being submitted to the slurm system and the job completing. This time includes the time the job took to run on the node which ranged from 2 to 6 hours. Also included was at least one maintenance day that occurred during one of the 698 job tests. During this time nearly 13k jobs were submitted with an average latency of 36.4hrs.

Figure 8 shows the results of 4 tests on Cori where a large group of jobs were submitted. Unlike swif2, which distributed the submission of jobs over time as files were transferred one by one to NERSC, these tests submitted all jobs simultaneously[3]. The input files had already been copied to the scratch disk.

---

[3] "simultaneously" here means a script was run on Cori to submit individual jobs one after
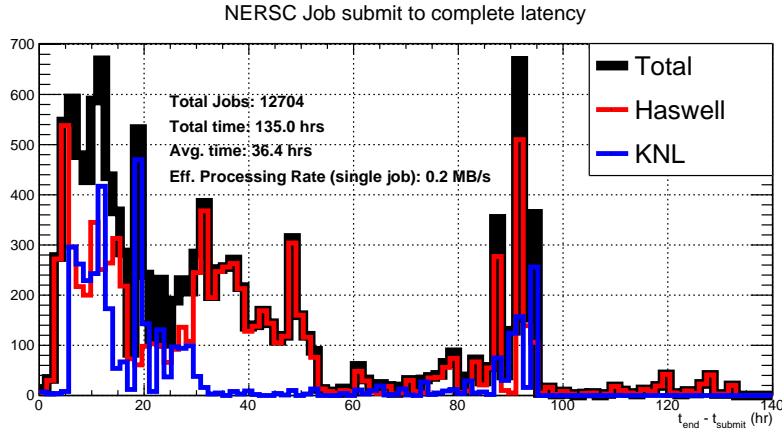
Figure 7: Latency between submission time and job completion time at NERSC. Several tests were done between July and October 2018 that are included here. Some latency can be attributed to maintenance days. Times also include the time for the job to run which was 2.4 hr on Haswell and 5.8 hr on KNL.

The intent of this exercise was to get an indication of how quickly slurm would dispatch the jobs without the added complication of file transfer latency. The plots show that the combined Cori I and Cori II availability to be 800-1000 simultaneous jobs, though there may be long periods when few jobs are dispatched. These tests spanned from 1 to 6 days so it is difficult to extrapolate to a campaign that is expected to span more than 1 month. It does seem to indicate though that running jobs on both Cori I and Cori II tends to give a significant benefit in effective throughput. At this time, this means we must split the jobs into 2 workflows in swif2 and manage them separately.

3   Anticipated Usage in AY2109

GlueX has accumulated 1.5PB of data in the Spring 2018 run and is currently at the start of the Fall 2018 run where another 1.0PB is anticipated. It is expected that 2 passes over the Spring data set and 1 pass over the Fall data set will be done offsite at NERSC during AY2019. Jobs from RunPeriod-2018-01 monitoring launch ver17 averaged 3.6 hrs on Cori I nodes which cost 288 NERSC units at 80 units per node hour. Benchmarking reconstruction of the same subset of 2018 data on both Cori I and Cori II indicates that Cori II takes 2.4 times longer to than on Cori I. Thus, at 96 NERSC units per node-hour on Cori II, jobs there will each cost 829 units. Table 1 summarizes the anticipated

---

another taking 10-20 minutes to submit all jobs. The submission rate can be seen as the slope of the far left edge of the blue curves.
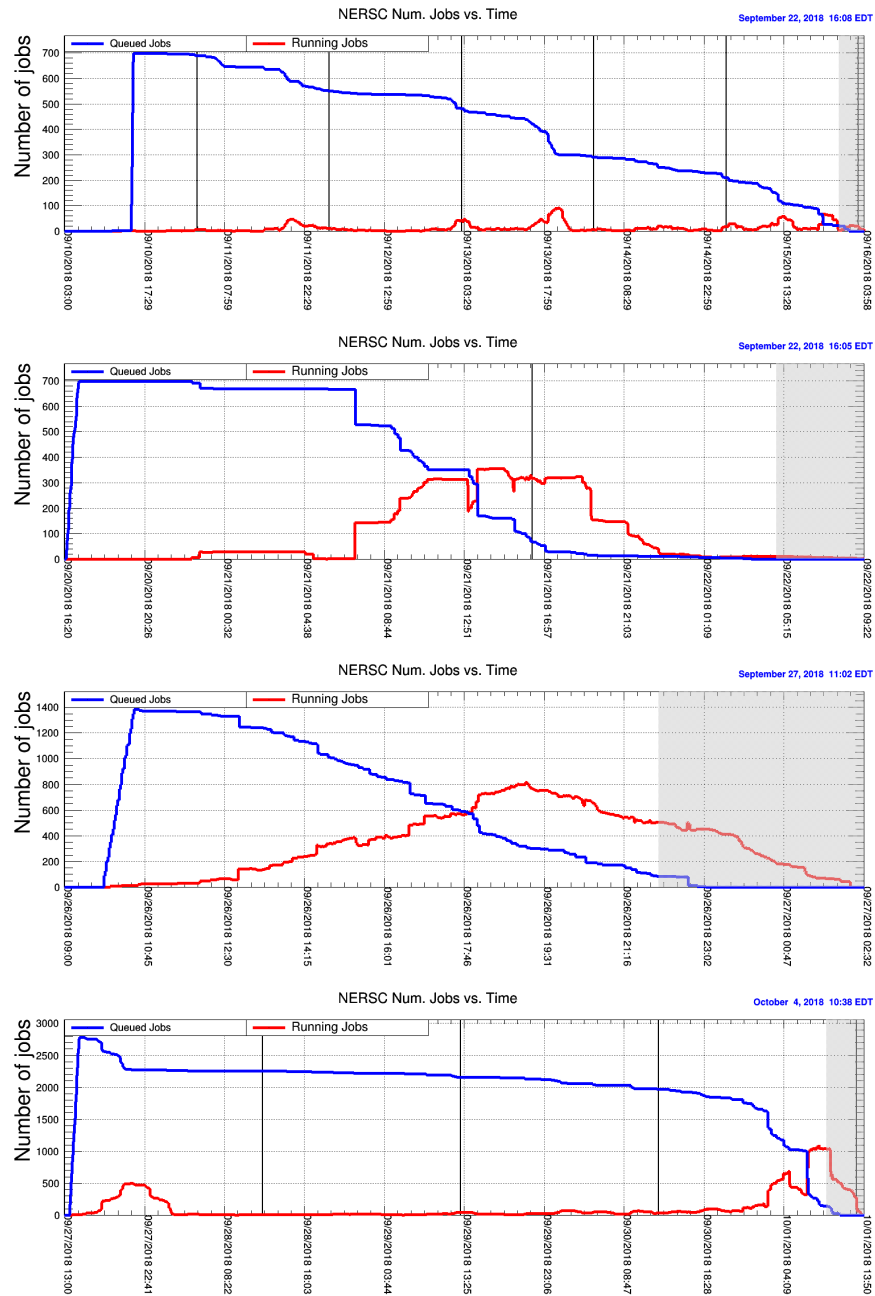
Figure 8: Instantaneous tally of number of queued (blue) and number of running(red) jobs on Cori for various tests. In order from top to bottom the plots are: 698 jobs Cori I only; 698 jobs Cori II only; 1496 jobs split between Cori I and Cori II; 2892 jobs split between Cori I and Cori II.

Total data to transfer to NERSC	4PB
Total jobs to be run at NERSC	200k
NERSC units per job Cori I	288
NERSC units per job Cori II	829
NERSC units required for Cori I	28.8M
NERSC units required for Cori II	82.9M
Total NERSC units required	**112M**

Table 1: Estimated NERSC units required by GlueX for AY2019.

NERSC usage requirement for AY2019.

3.1 Disk Space

Disk space usage is determined by the maximum number of nodes we may feed simultaneously at steady state. There are a total of 11k nodes in Cori which are typically in heavy demand. Haswell nodes can consume data at 20GB/3.6hr on average and KNL nodes at 20GB/8.6hr. The average rate of data consumption assuming all jobs are done on Cori I is therefore 1.5MB/s. If JLab is able to double their bandwidth to 20Gbps in the next few months[4] we may expect to get 10Gbps of usable bandwidth or roughly 1GB/s. This would be enough to feed 667 nodes. This means at least 667 x 20 GB files on disk for the live jobs and another 2 files for each job for queued files. This totals 40 TB of input raw data. Space for the output will be and additional 33% of that or 13.3 TB. All of these files will need only temporary storage at NERSC so are best suited for the scratch disk. A minimum of 53.3TB is therefore required. An additional 6TB would be desirable for use outside of production running (e.g. potential monitoring launch over data temporarily kept at NERSC). Thus, the total request will be for **60TB** of scratch disk space.

---

[4]anticipated