



NERSC SIGEUF

NERSC Special Interest Group on Experimental User Facilities

NUG Annual Meeting

Aug. 17, 2020



JOINT GENOME INSTITUTE

UNITED STATES DEPARTMENT OF ENERGY

2800 MITCHELL DR.

NERSC Special Interest Group on Experimental User Facilities

What:

Series of meetings focussed on a presentation by either a user or staff on issues with processing large experimental data sets at NERSC

Who:

Group of NERSC User's and Staff using NERSC to process data from external facilities

Why:

- Share issues and experiences that may benefit other users
- Get insight from staff regarding details of NERSC systems

When:

We met online most Wednesdays at 10am PST from Apri. 22 - Jul 15

NERSC Website:

<https://www.nersc.gov/users/NUG/sig-for-experimental-facility-users/>

The screenshot shows the NERSC website header with the logo and tagline "Powering Scientific Discovery Since 1974". A search bar and navigation links like "Site Map | My NERSC | Login | Share" are visible. The main navigation menu includes "HOME", "ABOUT", "COVID-19 SUPPORT", "SCIENCE", "SYSTEMS", "FOR USERS", "NEWS & PUBLICATIONS", "R & D", "EVENTS", and "LIVE STATUS". The "FOR USERS" section is expanded, listing various resources such as "Getting Help", "Getting Started", "Accounts & Allocations", "Documentation", "Policies", "My NERSC", "Job Logs & Statistics", "Training & Tutorials", "NERSC Users Group", "Monthly NUG Webinars", and "Annual Meetings". The "SIG Experimental Facility Users" link is highlighted in green. The main content area features the heading "SIG EXPERIMENTAL FACILITY USERS" in green, followed by a paragraph stating that NUG is sponsoring a Special Interest Group (SIG) within the NERSC Users Group for Experimental Facility Users. A second paragraph explains the group's purpose: to provide a forum for exchanging best practices, knowledge, and tools, and to provide feedback to NERSC staff on how to improve support for these workflows and policies. A final paragraph mentions the next meeting on Wednesday, May 6, 2020, at 10 am (Pacific time) and provides a link to the meeting archive: <https://www.jlab.org/indico/event/383/>.

NERSC
Powering **Scientific Discovery** Since 1974

Site Map | My NERSC | Login | Share
search... >>

HOME ABOUT COVID-19 SUPPORT SCIENCE SYSTEMS **FOR USERS** NEWS & PUBLICATIONS R & D EVENTS LIVE STATUS

FOR USERS

- ▶ Getting Help
- ▶ Getting Started
- ▶ Accounts & Allocations
- ▶ Documentation
- ▶ Policies
- ▶ My NERSC
- ▶ Job Logs & Statistics
- ▶ Training & Tutorials
- ▶ **NERSC Users Group**
- Monthly NUG Webinars
- Annual Meetings
- SIG Experimental Facility Users**

Home » For Users » NERSC Users Group » SIG Experimental Facility Users

SIG EXPERIMENTAL FACILITY USERS

NUG is sponsoring a Special Interest Group (SIG) within the NERSC Users Group for Experimental Facility Users.

The group was formed so that NERSC users that process data from experimental and observational facilities sponsored by the Department of Energy Office of Science have forum for exchanging best practices, knowledge and tools that are broadly applicable. The group will also provide feedback to NERSC staff on how to improve support for these workflows and provide input on policies that effect this growing workload at NERSC.

The next meeting will be held on Wednesday, May 6, 2020, at 10 am (Pacific time). For more information, including how to connect to the meeting an archive of all resources from previous meetings, please see <https://www.jlab.org/indico/event/383/>.

Inidco Website:

<https://www.jlab.org/indico/event/383/>

(all talks and videos including discussions are archived)

NERSC Users Group Special Interest Group on Experimental Facilities

chaired by David Lawrence (Jefferson Lab)

from Wednesday, 22 April 2020 at **10:00** to Wednesday, 16 December 2020 at **11:20** (US/Pacific)

Manage ▾

Description **Connection information:**

Meeting URL <https://bluejeans.com/858265685>
Meeting ID **858 265 685**


Want to dial in from a phone?
Dial one of the following numbers:
+1.888.240.2560 (US Toll Free)
(see all numbers - <https://www.bluejeans.com/premium-numbers>)

Enter the meeting ID and passcode followed by #

Connecting from a room system?
Dial: bjn.vc or 199.48.152.152 and enter your meeting ID & passcode

Go to day ▾

Wednesday, 22 April 2020

10:00 - 11:20	NUG SIG Experimental User Facilities	▾
10:00	Welcome 5' Speaker: David Lawrence (Jefferson Lab)	▾
10:05	Supporting our data workload at NERSC 10' Speaker: Sudip Dosanjh	▾
10:15	Best practices and tips for users from experimental facilities running at NERSC 30' Speaker: Bjoern Enders Material: Slides 	▾

< this page intentionally left blank >

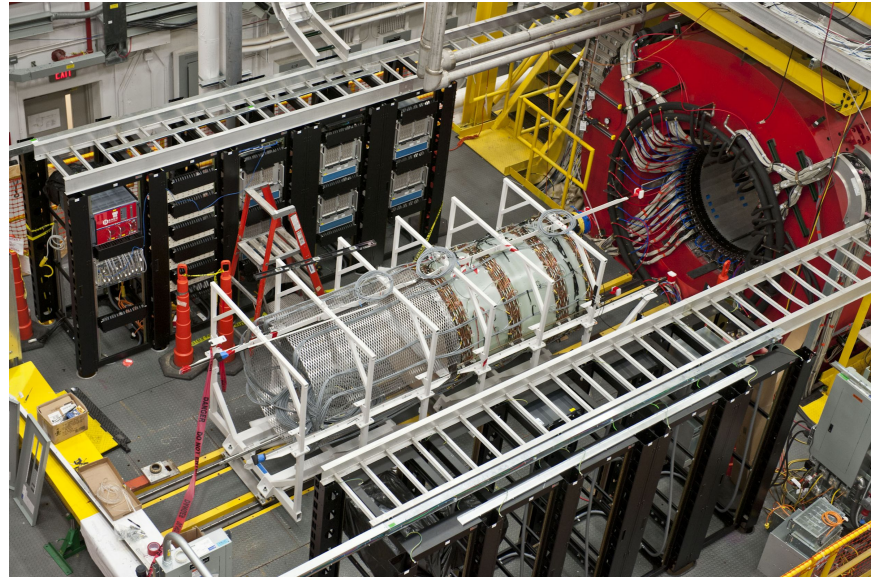
Offsite Data Processing for the Experiment

David Lawrence - JLab
on behalf of the
GlueX Collaboration

~~Thursday November 7, 2019
CHEP2019, Adelaide, Australia~~

~~Wed. May 6, 2020
NUG Special Interest Group on
Experimental User Facilities~~

Jefferson Lab



Forward Drift Chamber installation in GlueX Dec. 2013

Mon. Aug. 17, 2020
NUG Meeting: Breakout session on
Experimental User Facilities

Aerial photo taken Apr

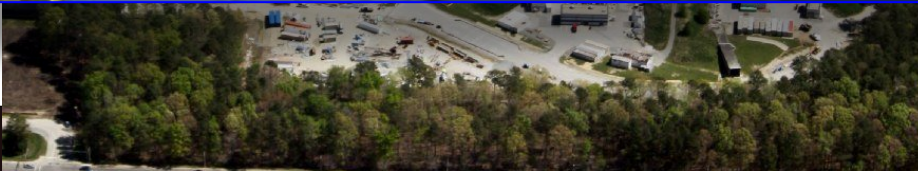
Hall-D
Complex

Electron beam

- continuous (250MHz, 4 structure in
- Polarized e
- Upgraded t (from 6GeV
- 70 μ A max (200 μ A max @ 6GeV)



**Thomas Jefferson National Accelerator Facility (JLab)
Newport News, Virginia, USA**



Aerial photo

Hall-D Complex

Electron beam accelerator

- continuous-wave
(250MHz, 4ns bunch structure in halls)
- Polarized electron beam
- Upgraded to 12GeV
(from 6GeV)
- 70 μA max @ 12Gev
(200 μA max @ 6GeV)



GlueX Computing Numbers

Data Volume

	2017 (low intensity GlueX)	2018 (low intensity GlueX)	2019 (PrimEx+ high intensity GlueX)	2020 (high intensity GlueX)
actual (raw data only)	0.91PB	3.11PB	0.40PB*	
model (raw data only)	0.86PB	3.17PB	1.56PB	6.06PB
actual (production)	1.26PB	1.21PB*	0.62PB*	
model(production)	0.61PB	3.08PB	1.94PB	4.34PB
Total Data (actual)	2.17PB	4.32PB*	1.02PB*	
Total Data (model)	1.47PB	6.25PB	3.5PB	10.4PB

CPU (Haswell core equivalent from model)

	2017 (low intensity GlueX)	2018 (low intensity GlueX)	2019 (PrimEx)	2019 (high intensity GlueX)
Real Data CPU	21.3Mhr	67.2Mhr	6.4Mhr	39.6Mhr
MC CPU	3.0Mhr	11.3Mhr	1.2Mhr	8.0Mhr
Total CPU	24.3Mhr	78.4Mhr	7.6Mhr	47.5Mhr

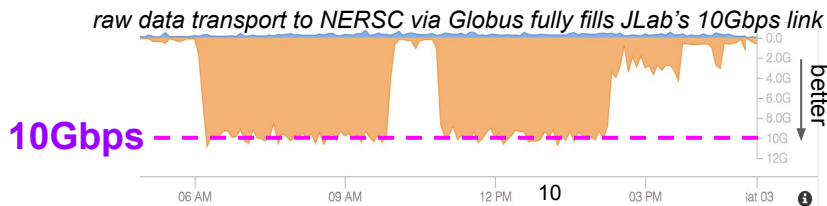
GlueX Offsite Computing Model

OSG, NERSC, PSC jobs use the same:

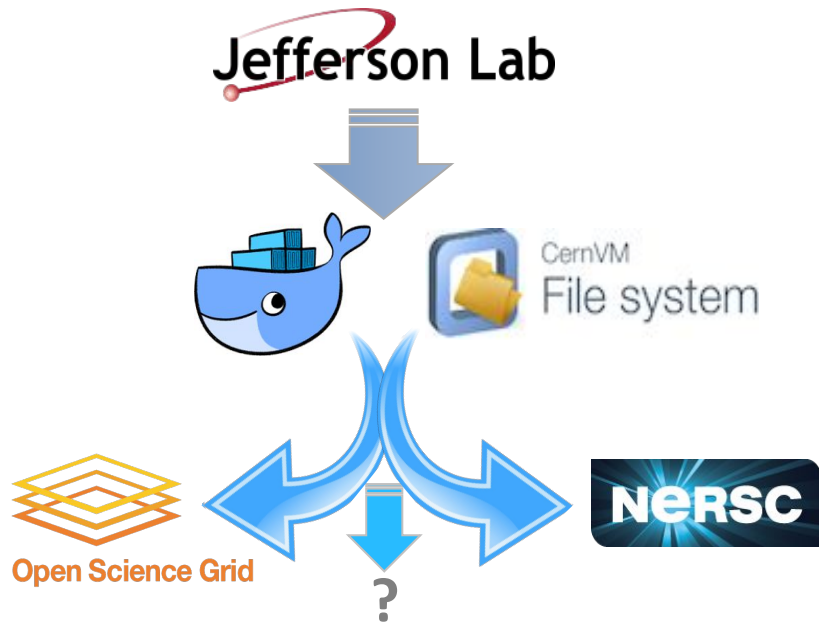
- Docker container (*converted to Singularity and Shifter*)
 - same container used for **Cori I** (*Haswell*) and **Cori II** (*KNL*)
- CVMFS share
 - GlueX Software builds for CentOS 7
 - 3rd party software (e.g. ROOT)
 - Calibration Constants (CCDB SQLite file)
 - Resource files (field and material maps)

Data Transport:

- NERSC and PSC: Globus
- OSG: Condor



Offsite Data Processing for the GlueX Experiment - D. Lawrence(JLab) CHEP19, Adelaide, AU

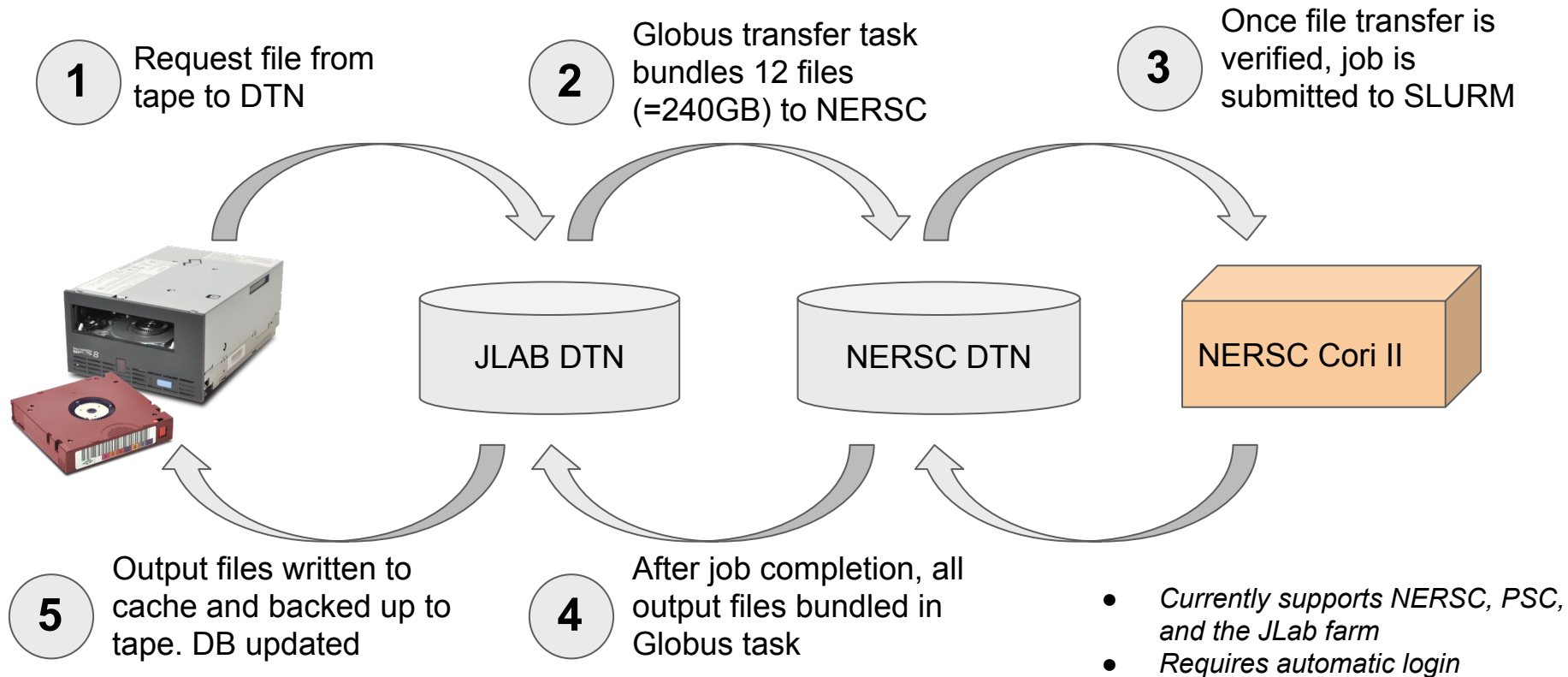


Containerized software runs at NERSC on both **Cori I** (*Haswell*) and **Cori II** (*KNL*)

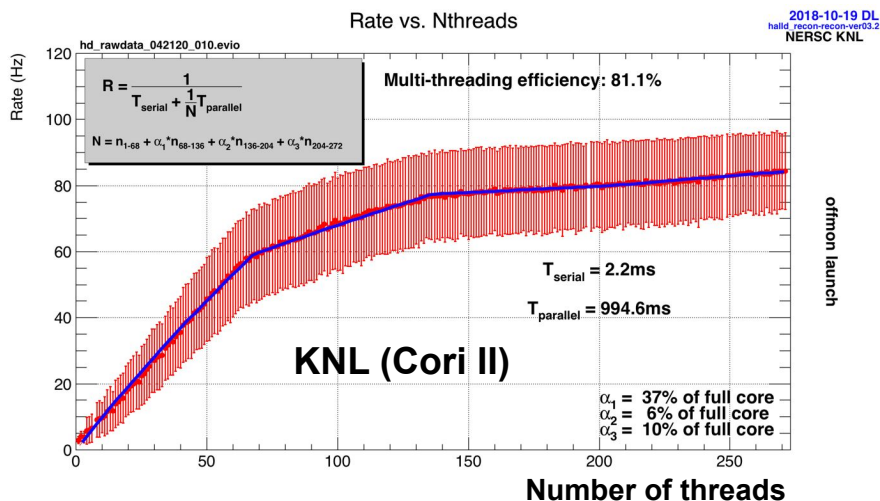
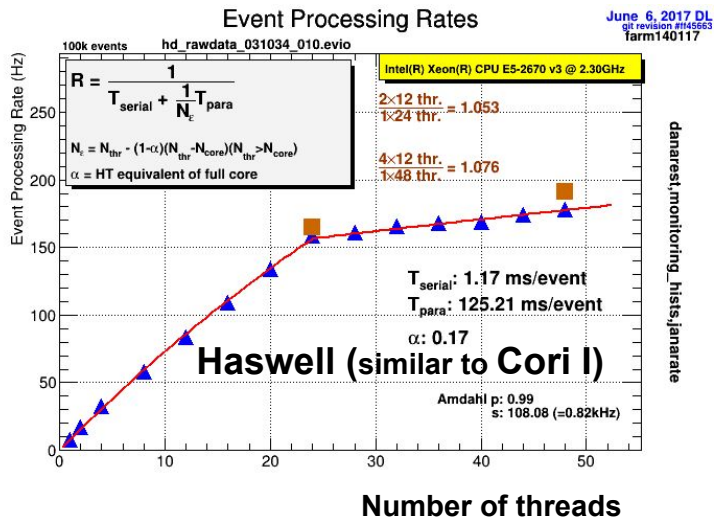
OSG = Open Science Grid
NERSC = National Energy Research Scientific Computing Center
PSC = Pittsburgh Supercomputing Center

SWIF2 - Job workflow tool

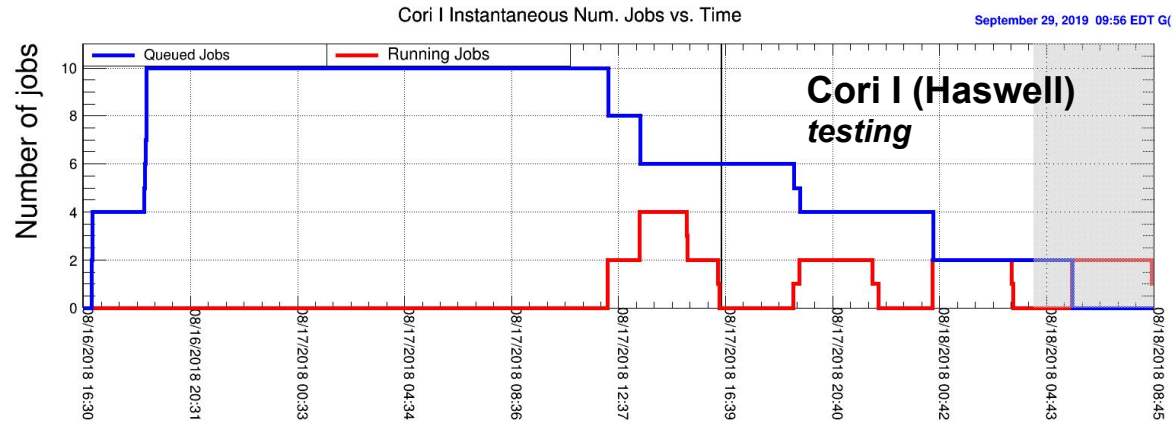
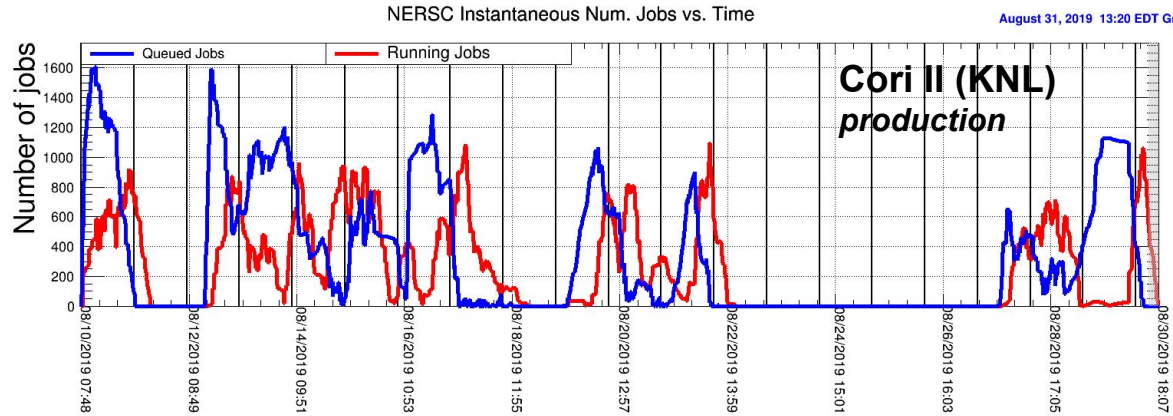
Manage file transfer and job submission through limited disk resources



GlueX Allocation AY2019	58.5M NERSC Units
Input file size	20GB (91.9k jobs so far in 2019)
Wall Time/file on Cori I (Haswell)	3 hours
Wall Time/file on Cori II (KNL)	6.9 hours



GlueX @ NERSC - Backfilling



Job scheduler at NERSC is extremely poorly matched to our job shape:

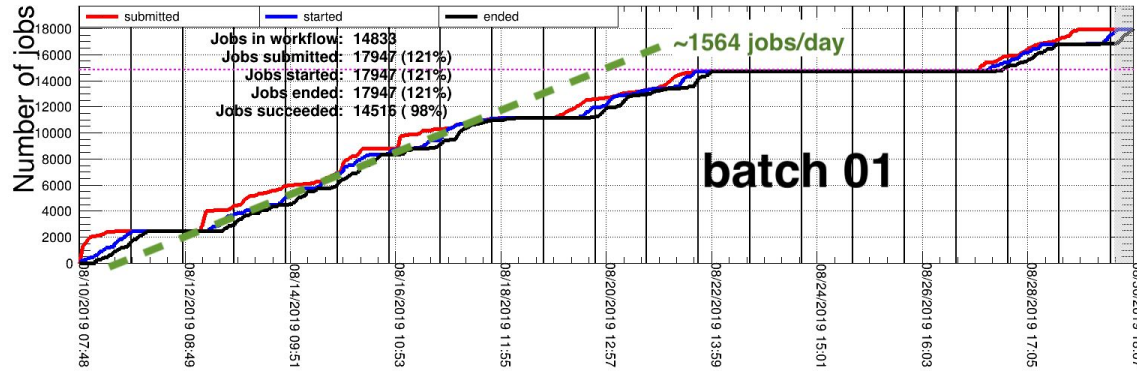
- Schedule at most 2 “jobs” at a time via priority and all others must schedule via backfill
- Scheduler ignores number of nodes and time requested when determining priority
- 64 nodes x 48 hours = 1 node x 3 hours

Suspect most of our jobs run via “backfill” since they are small and fit in cracks. (*Test on Cori I supports this*)

“Regular” vs. “Low” priority queue on Cori II

NERSC Integrated Num. Jobs vs. Time

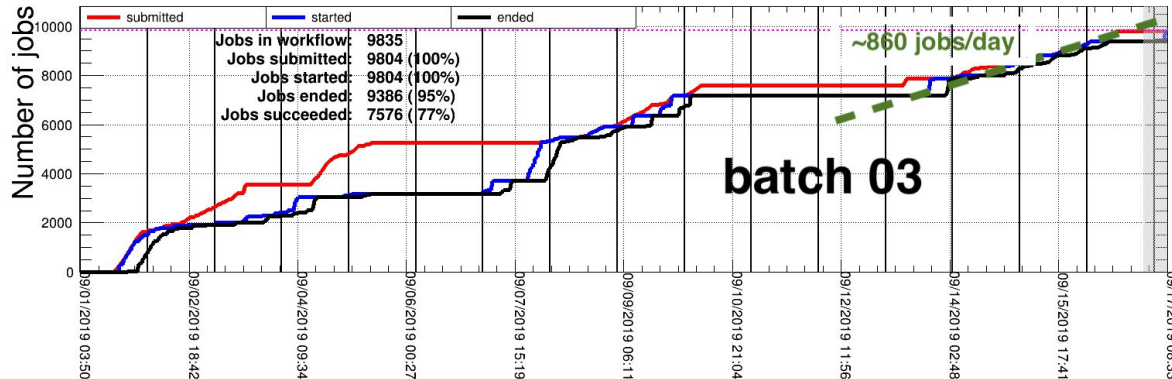
August 31, 2019 13:21 EDT G



“normal” queue on Cori II
Aug. 13 - Aug. 18, 2019
~1564 jobs/day

NERSC Integrated Num. Jobs vs. Time

September 17, 2019 08:34 EDT G



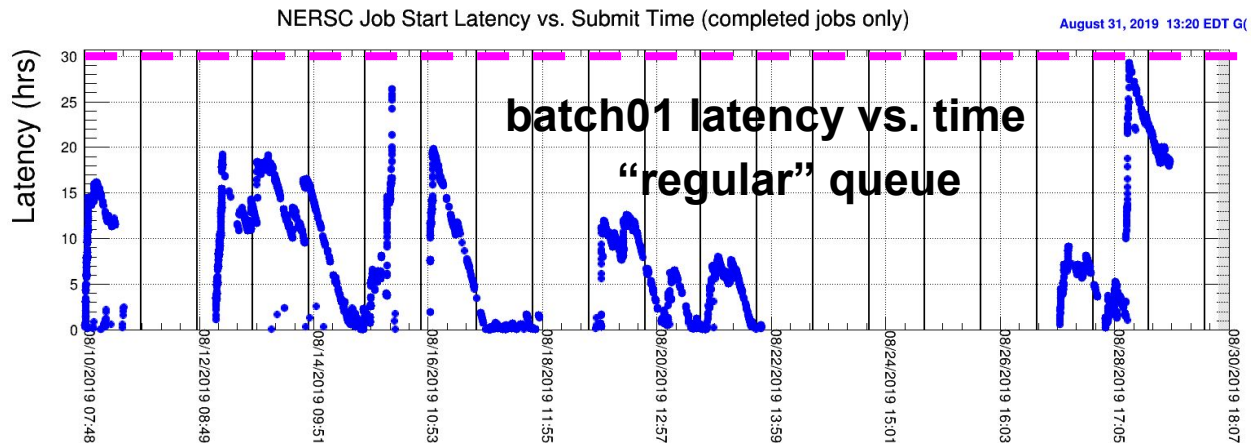
“low” queue on Cori II
Sep. 14 - Sep. 16, 2019
~860 jobs/day

Summary and Outlook

- GlueX is now able to reconstruct large Experimental Nuclear Physics data sets offsite
 - NERSC, PSC
 - Lightweight container used for all offsite HPC computing
 - Software distributed via CVMFS
 - SWIF2 manages workflow
- NERSC
 - Scheduler poorly matched to our natural job size
 - Backfilling saves us (and benefits them!)
 - Job rate fluctuates but averages ~1k/day (=20TB/day) *no longer true in 2020. rate has dropped to closer to 1k/week*
- PSC
 - Better matched to our natural job size but smaller resource
 - Job rate fairly steady ~0.3k/day (=6TB/day)
- Most simulation is being done on OSG

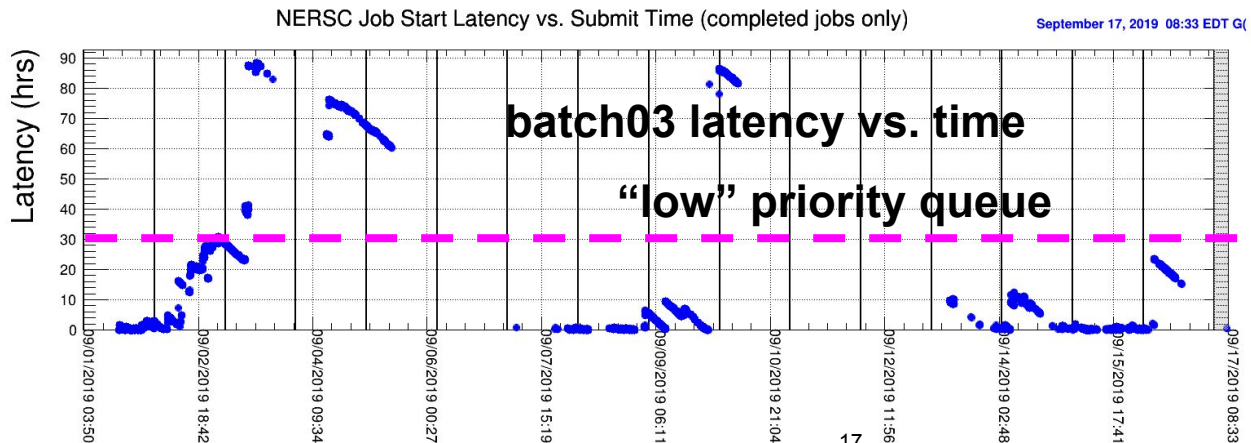


NERSC Job Start Latency



Time between submitting job to slurm on NERSC Cori II and job starting

Everything is anecdotal!

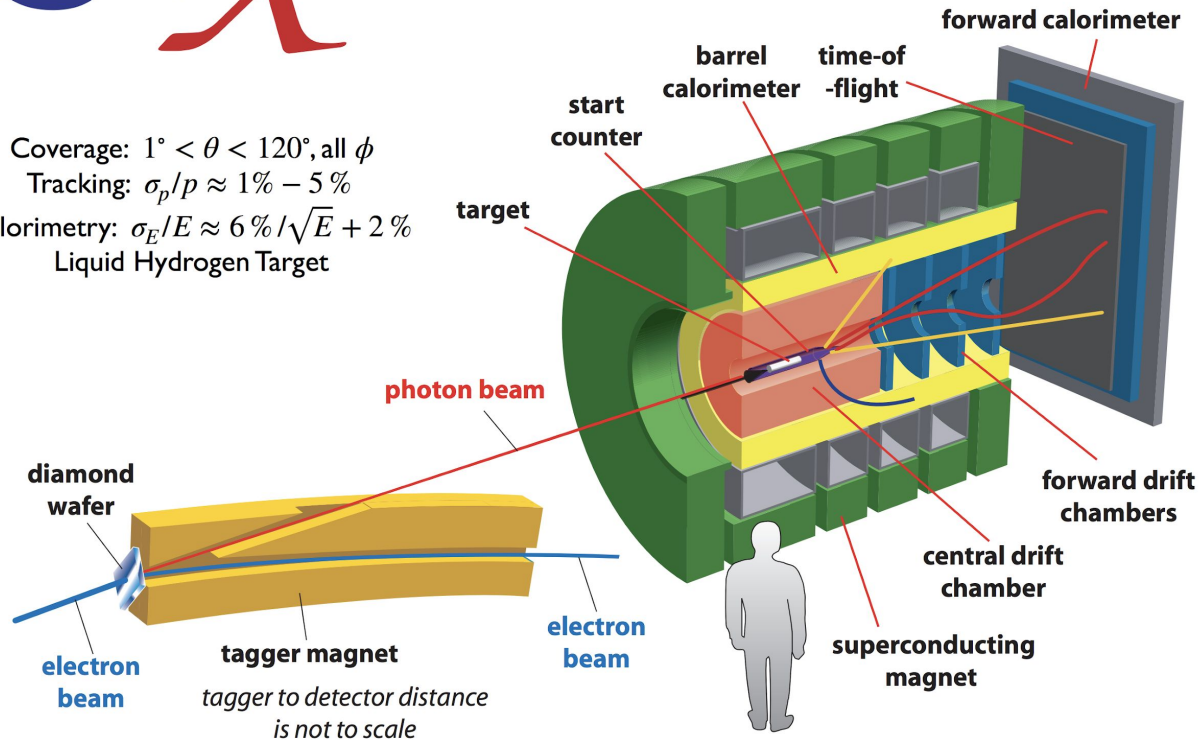


The GlueX Detector

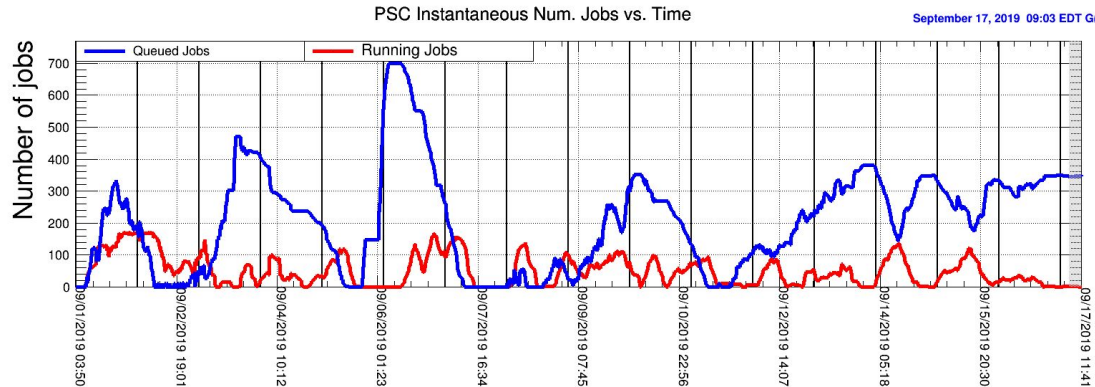
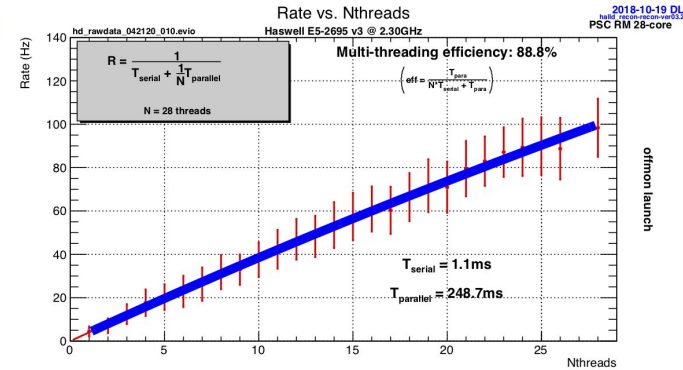
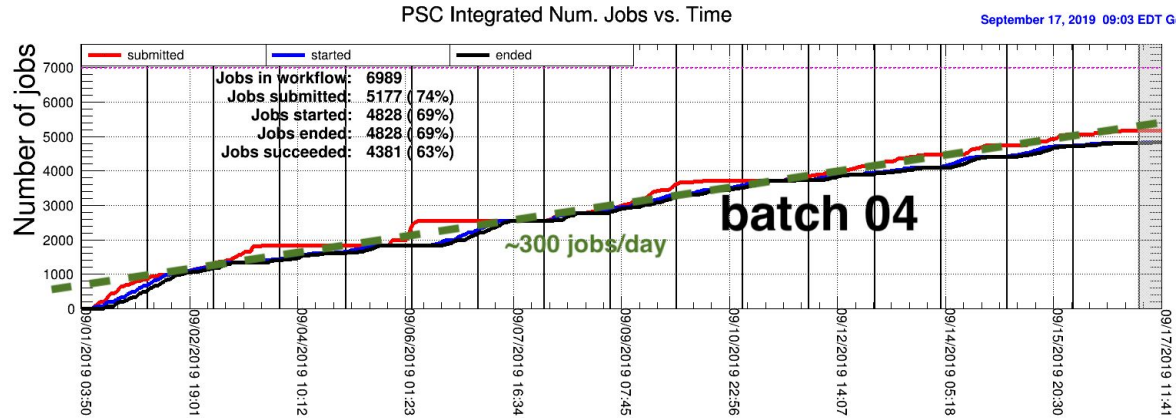
- Large Superconducting Solenoid
- Fixed target (30cm LH2)
- Coherent bremsstrahlung polarized photon source
- 38k Detector Channels
- Charged particle tracking, Segmented Calorimetry, PID



Coverage: $1^\circ < \theta < 120^\circ$, all ϕ
Tracking: $\sigma_p/p \approx 1\% - 5\%$
Calorimetry: $\sigma_E/E \approx 6\%/\sqrt{E} + 2\%$
Liquid Hydrogen Target



GlueX @ Pittsburgh Supercomputing Center (XSEDE)



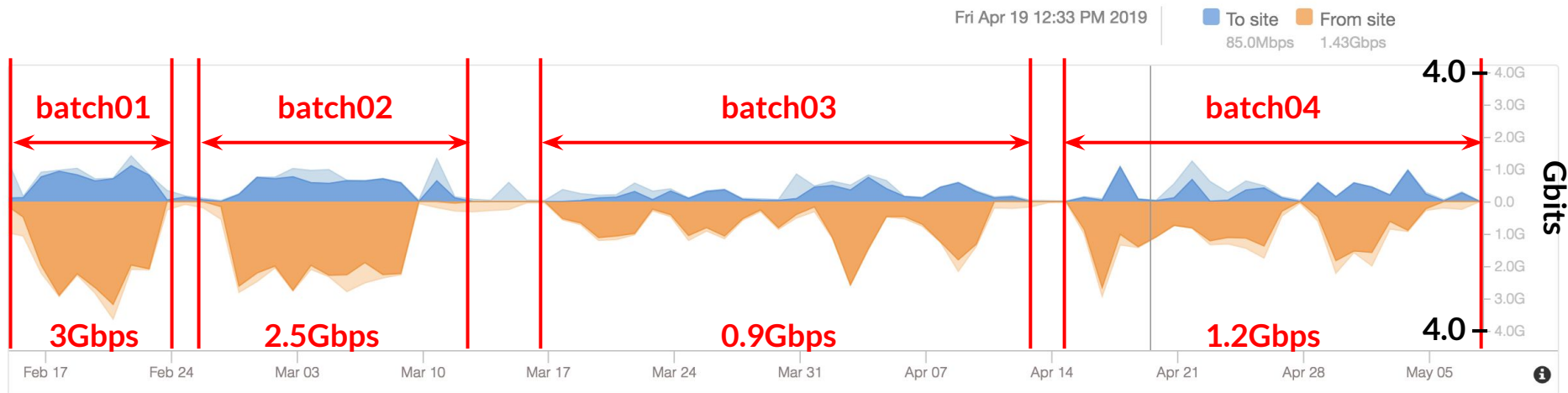
PSC Bridges:

- 28 cores/node (no HT)
- 4.2 hours/job
- 6,989 jobs

Smaller than NERSC, but more steady and smaller failure rate

ESNet data transfer rates to/from NERSC

- Currently have 10Gbit connection
- Will activate second 10Gbit connection this summer
- Proposed 100Gbit upgrade in 2020 or 2021



- Anti-correlation observed between transfer rate and Lustre usage
- Test done using OSG16 node, disk speed an issue (longer story, ask Thomas)
- New DTN (Data Transfer Node) being configured with SSD disks for test
- Currently: 10% of files go through OSG node and 90% via cache(=Lustre)

Overview of Jefferson Lab

- Department of Energy National Laboratory with research mission in Nuclear Physics
- In operation since 1995
- Managed for DOE by Jefferson Science Associates, LLC
 - Joint venture of Southeastern Universities Research Association and PAE
- Our primary research tool is CEBAF (Continuous Electron Beam Accelerator Facility) – unique in the world



Jefferson Lab by the numbers:

- 700 employees
- FY2018 Budget: \$162.4M
- 169 acre site
- 1,600 Active “User Scientists”
- 27 Joint faculty
- 608 PhDs granted to-date (211 in progress)
- K-12 programs serve more than 13,000 students and 300 teachers annually