

LQCD-ext II Project Report

Bill Boroski
LQCD-ext II Project Manager
boroski@fnal.gov

USQCD All-Hands Meeting
Brookhaven National Lab
April 26-27, 2019

Outline

- ▶ LQCD–ext II progress updates
- ▶ Organizational updates
- ▶ User survey results and plans
- ▶ Web site and documentation updates
- ▶ 5–year extension planning (FY20–24)

LQCD-ext II Progress Update

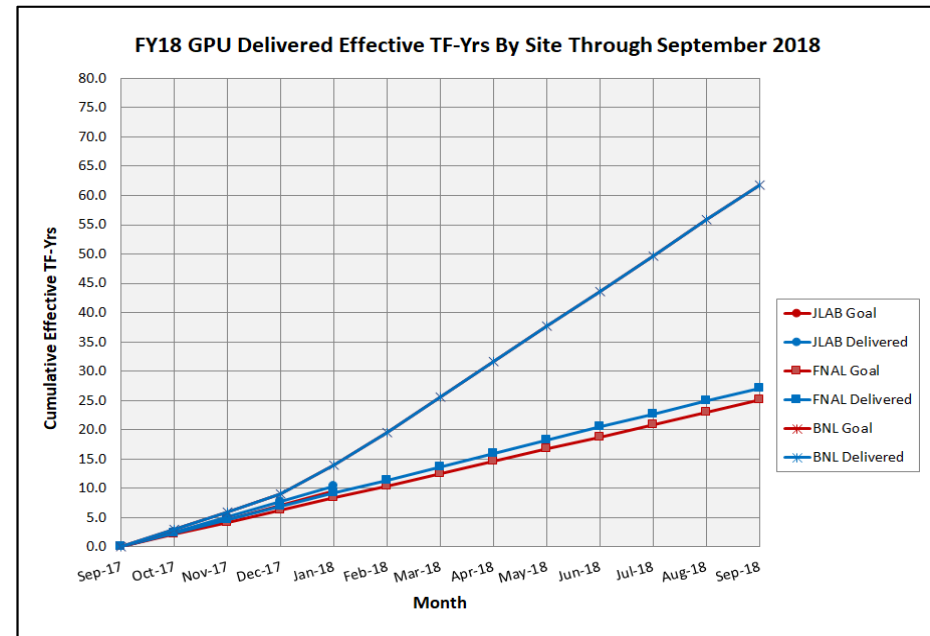
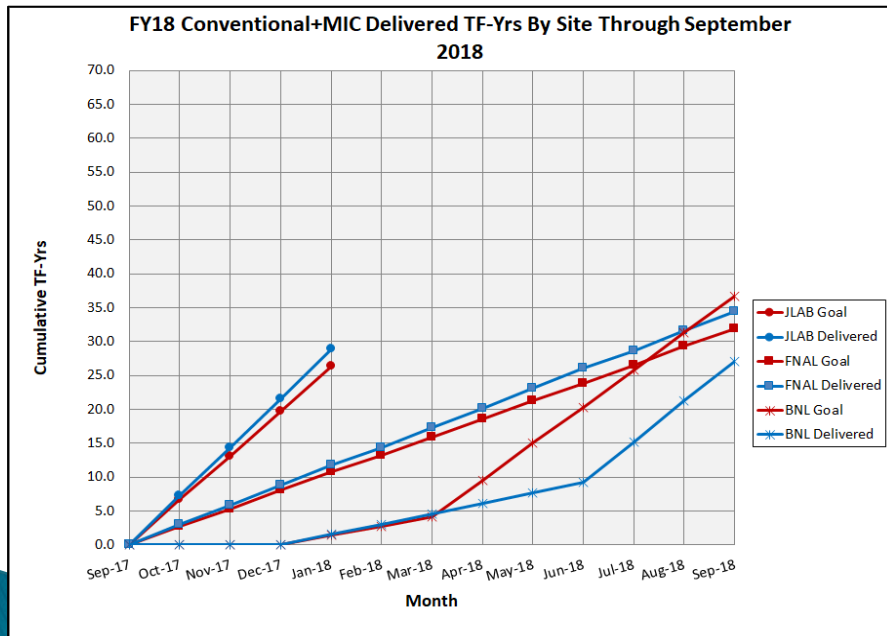
- ▶ We're in the fifth year of the current 5-year project (Oct 2014–Sep 2019)
- ▶ Operations continue to run smoothly at all of our host sites
- ▶ In January 2018, DOE Office of Nuclear Physics (NP) announced its intent to establish an NP-funded dedicated hardware project at JLab (*Nuclear and Particle Physics LQCD Computing Initiative, or NPPLCI*). Operations under the new structure began in earnest in February 2018. *(NPPLCI details will be discussed in another presentation later today)*
- ▶ LQCD-ext II project currently consists of deployments and operations at BNL and FNAL and is funded entirely by DOE Office of High Energy Physics (HEP)
- ▶ We continue to maintain information-sharing and knowledge transfer between the three host sites (bi-weekly Site Manager Meetings and quarterly DOE calls). Resources from all sites are available to all of USQCD.
- ▶ Between LQCD-ext II and the new NP initiative at JLab, we have received the full \$14M of planned funding in accordance with the approved baseline plan, dated Oct 1, 2014

Entity	FY15 (HEP & NP)	FY16 (HEP & NP)	FY17 (HEP & NP)	FY18	FY19	Total
LQCD-ext II	\$2M	\$3M	\$3M	\$2M (HEP)	\$2M (HEP)	\$12M
NPPLCI	---	---	---	\$1M (NP)	\$1M (NP)	\$2M

LQCD-ext II Results - FY18

Delivered Computing:

	TF-yrs Delivered			Node-hrs Delivered
	Goal	Actual	% of Goal	Actual
Conventional Resources	95.0	90.3	95%	4,842,249
Accelerated Resources	96.4	99.3	103%	981,276



LQCD-ext II Progress: FY19 Year-to-Date

► FY19 (thru March 2019)

- BNL has brought online 600 TB of new tape storage and expanded the size of their IC machine by 54 nodes, from 162 to 216 nodes
 - Current configuration: 50% K80s, 50% P100s.
- FNAL is working on bringing online a new cluster and additional storage; will be online in early July, maybe sooner.

Delivered Computing (TF-yrs)							
	FY19 (Oct '18 thru Mar '19)			Cumulative (Oct '14 thru Mar '19)			% of Goal
	Goal	Actual	% of Goal	Goal	Actual	% of Goal	
Conventional Resources ²	22.1	22.8	103%	443.5	466.7	105%	
Accelerated Resources ³	27.5	28.0	102%	377.6	394.4	104%	

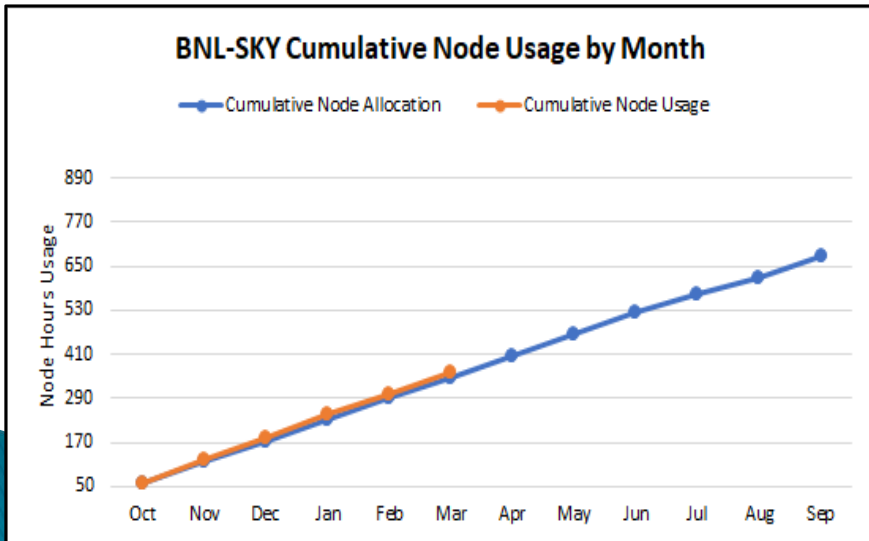
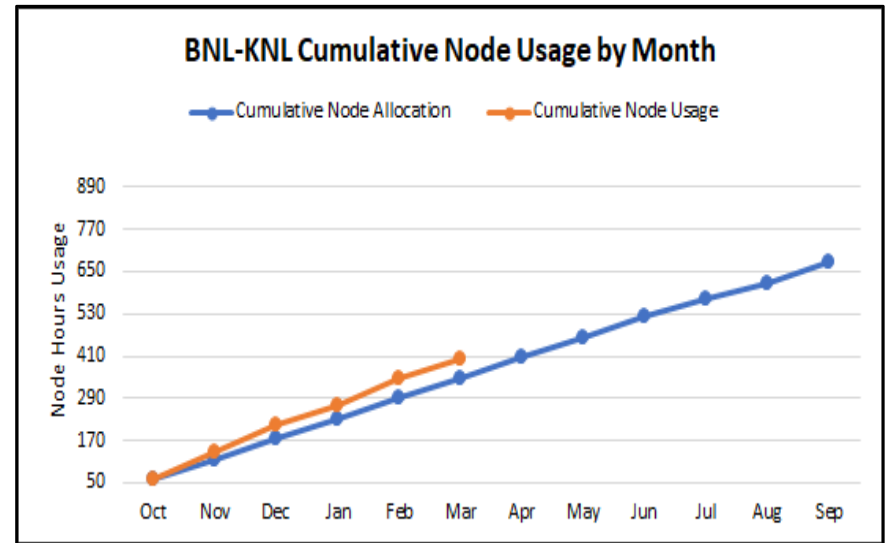
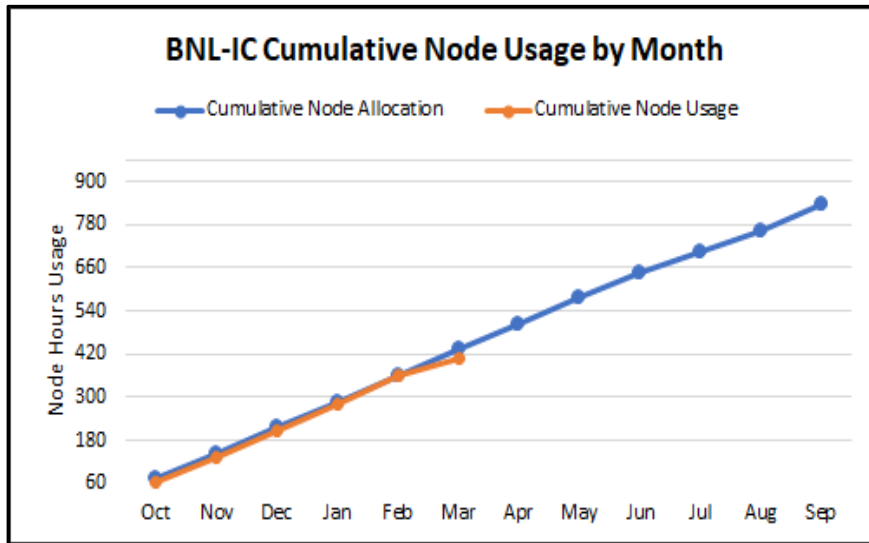
1) Conventional resources operational in FY19: Pi0, BNL-KNL, BNL-SKY.

DNR Mode (Bc, Ds)

2) Accelerated resources operational in FY19: Pi0g, BNL-IC.

DNR Mode (Dsg)

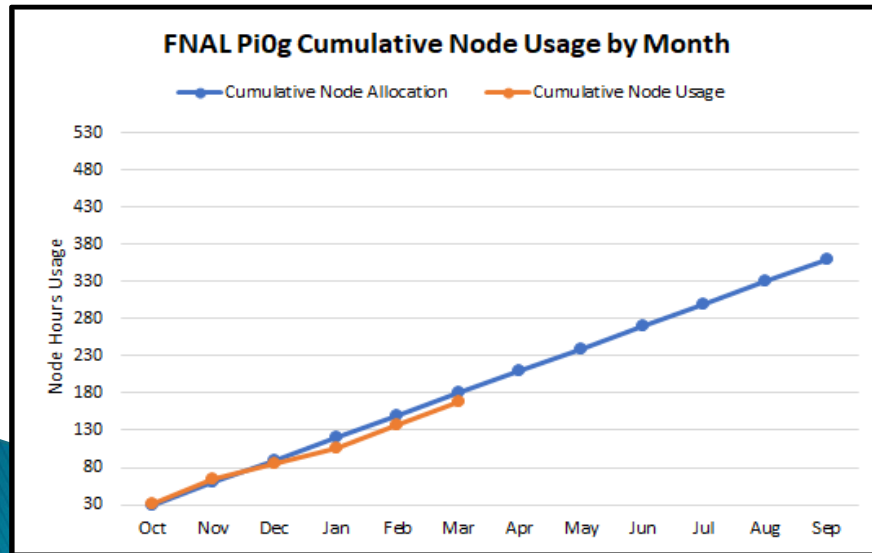
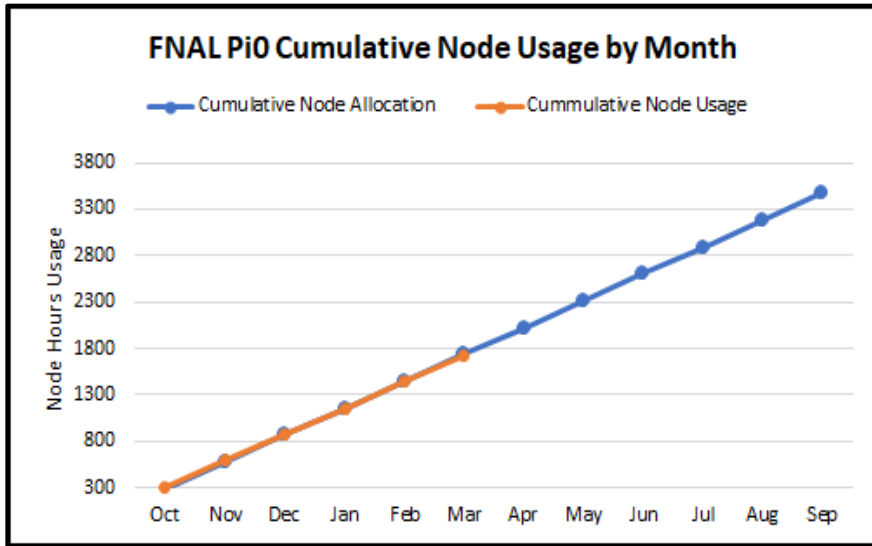
Utilization of our FY19 Allocations – BNL



FY19: Oct thru March:

System	Node Allocation	% of Cumulative Allocation Used
BNL-IC	72	94%
BNL-KNL	58	115%
BNL-SKY	58	103%

Utilization of our FY19 Allocations – FNAL



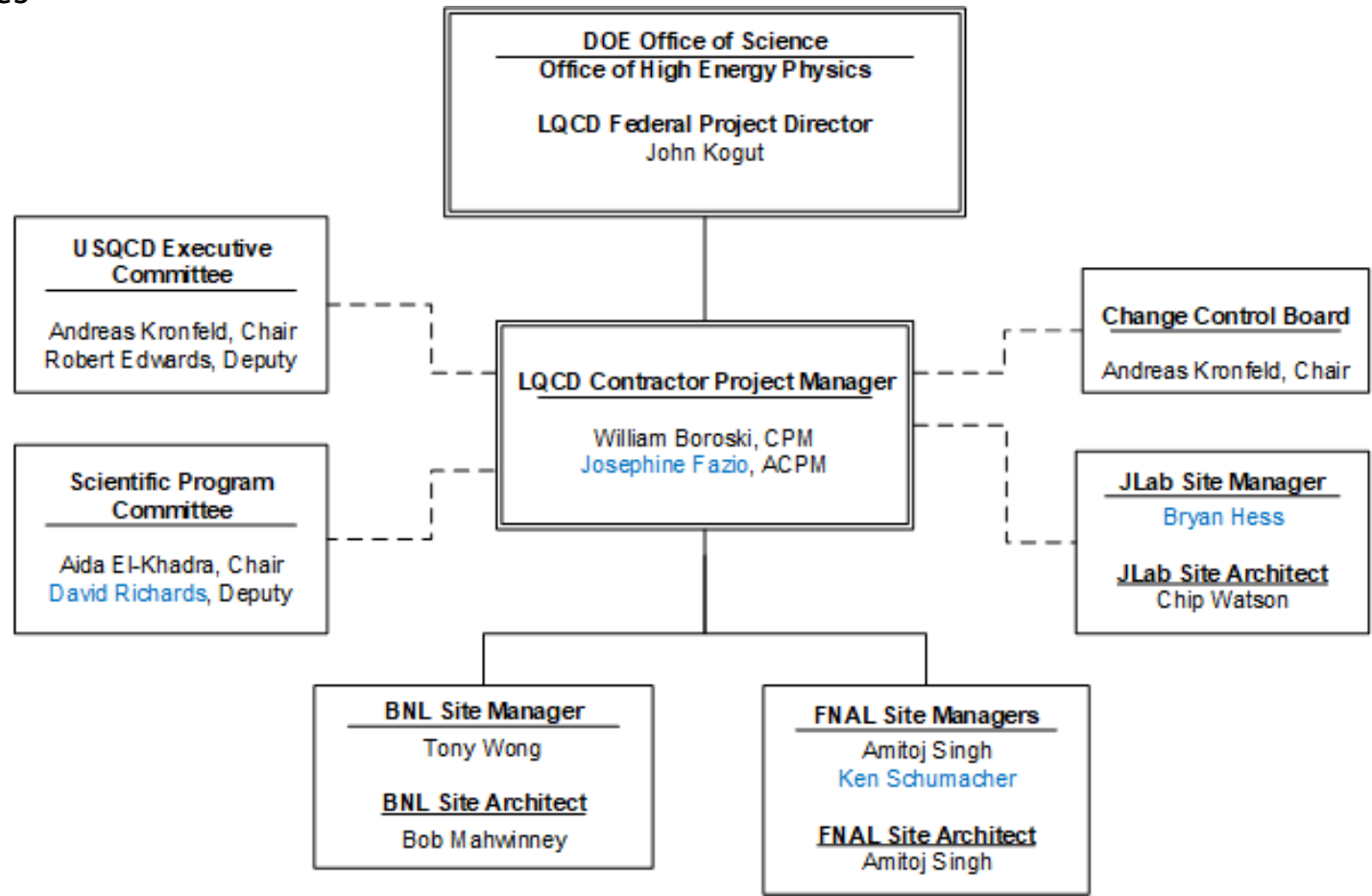
FY19: Oct thru March:

System	Node Allocation	% of Cumulative Allocation Used
FNAL-Pi0	289	100%
FNAL-Pi0g	30	93%

LQCD-ext II Integrated Project Team (IPT)

Organizational changes since last year:

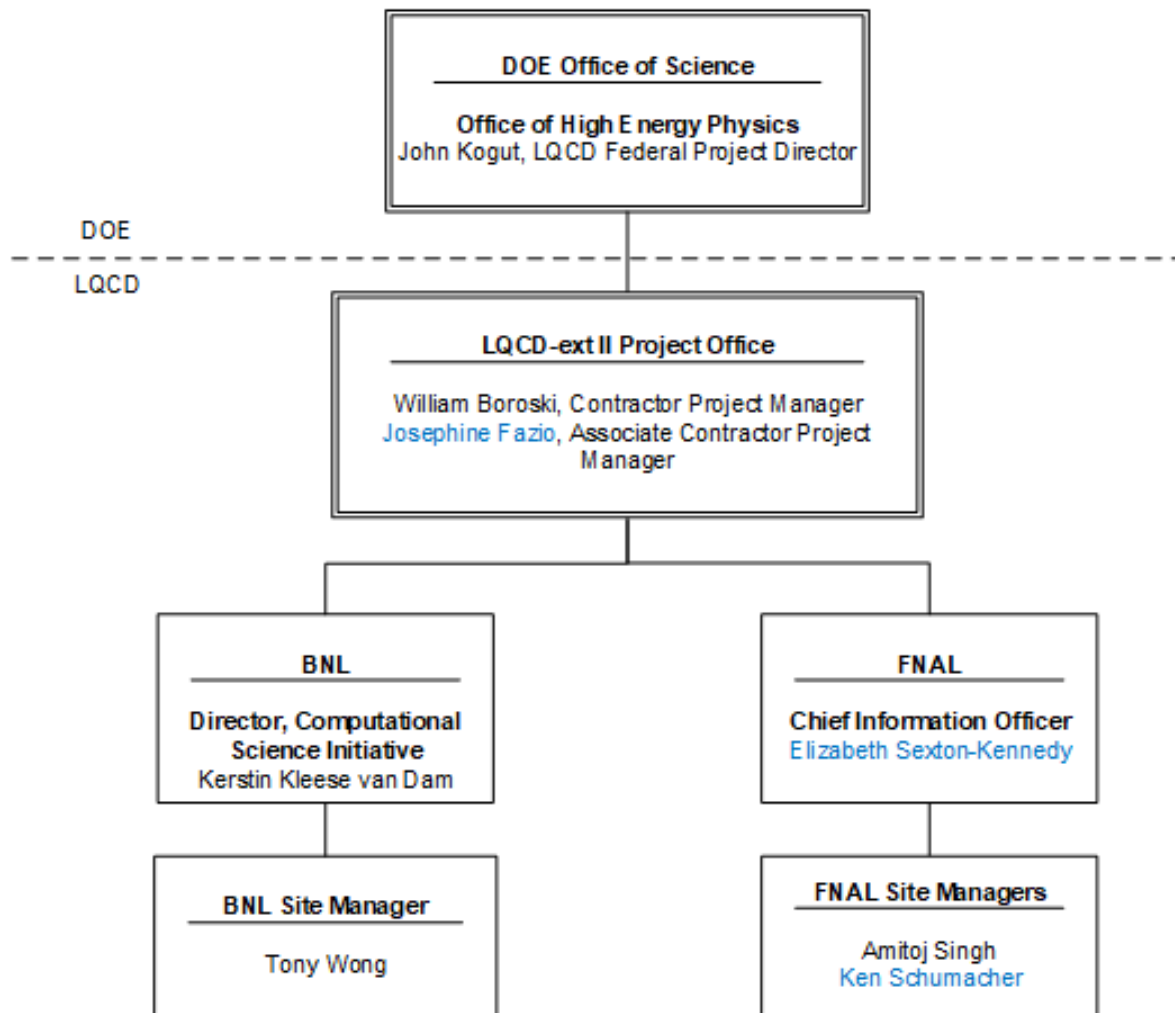
- Jo Fazio
- David Richards
- Ken Schumacher
- Bryan Hess



LQCD-ext II Integrated Management Team

Organizational changes since last year:

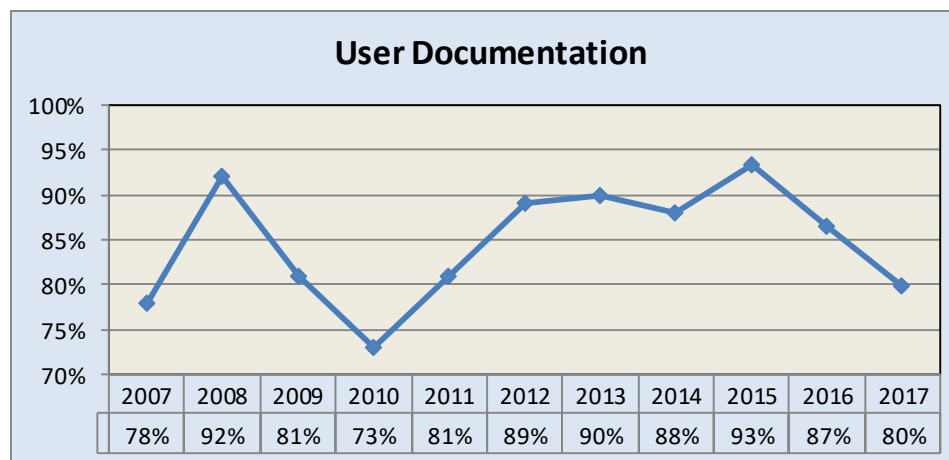
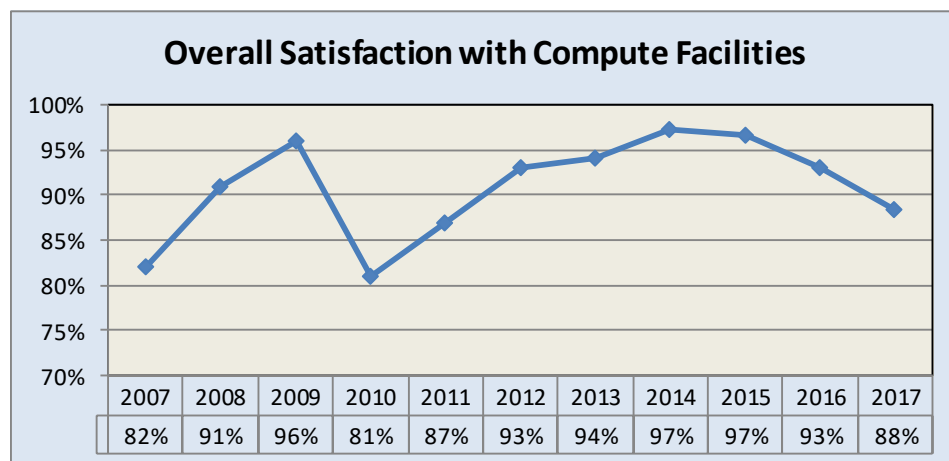
- Liz Sexton-Kennedy
- Jo Fazio
- Ken Schumacher



User Survey Results: Compute Facilities

- ▶ Thank you to everyone who participated in last year's User Survey.
- ▶ Total responses received=57; down from 73 in previous year

- ▶ Overall Satisfaction: 88%
 - FNAL:97%; other sites lower
- ▶ Problem areas
 - User documentation
 - Ease of access
 - User support
 - System reliability
- ▶ Significant effort to address documentation and user support shortcomings
 - Implemented documentation updates at FNAL
 - Created new documentation and implemented new processes at BNL.



Web Site Updates

- ▶ LQCD Project website available through USQCD.org or directly <https://www.usqcd.org/lqcd/WBS/>
 - Fixed some incorrectly routed links
 - Updated FNAL landing page
- ▶ New BNL web pages <https://www.usqcd.org/bnl/>
 - Created USQCD introduction web pages
 - Developed overview to help new users get started
 - Improved online system documentation

USQCD US Lattice Quantum Chromodynamics

USQCD home Physics program Software Hardware USQCD Collaboration Links and resources

Particle and Nuclear Physics

Installations BNL LQCD
Clusters FNAL LQCD
LQCD Project JLAB LQCD

USQCD is a collaboration of US scientists developing and using large-scale computers for calculations in lattice quantum chromodynamics.

Lattice QCD calculations allow us to understand the results of particle and nuclear physics experiments in terms of QCD, the theory of quarks and gluons.

BNL Computing Information for USQCD Users

Home

Site Map

Overview of BNL Clusters for USQCD Users

- Getting Help
- User Accounts and Access
- Available Clusters
- Filesystem Details
- File Transfers - Globus online
- Monitoring Usage

Overview of BNL Clusters for USQCD Users

The USQCD Collaboration contracts for the use of time on various clusters at BNL. These web pages contain information about the clusters of specific relevance to USQCD users. BNL maintains a general web page, <https://www.sdcc.bnl.gov/>, for the clusters, which are run by the Scientific Data and Computing Center (SDCC) at BNL. This SDCC page contains links to more information about the general configuration of the clusters, access to them and their attached storage.

USQCD time on the BNL clusters is allocated by the USQCD Scientific Program Committee (SPC). The USQCD liaison to BNL acts as an intermediary between the SPC and the BNL staff who manage the clusters. The liaison asks BNL staff to set up the accounts for each PI, and their collaborators, manages disk quotas for each project and monitors computer usage by each project. This usage is reported to USQCD project managers every month.

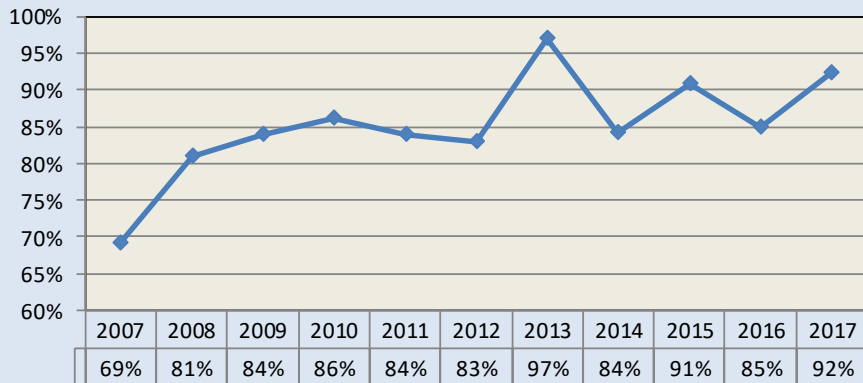
BNL staff manage the computing resources, including compute systems, storage systems (disk and tape), networks and queues. The installation and maintenance of software for general use is handled by BNL staff.

USQCD is one of a number of groups who use the BNL clusters.

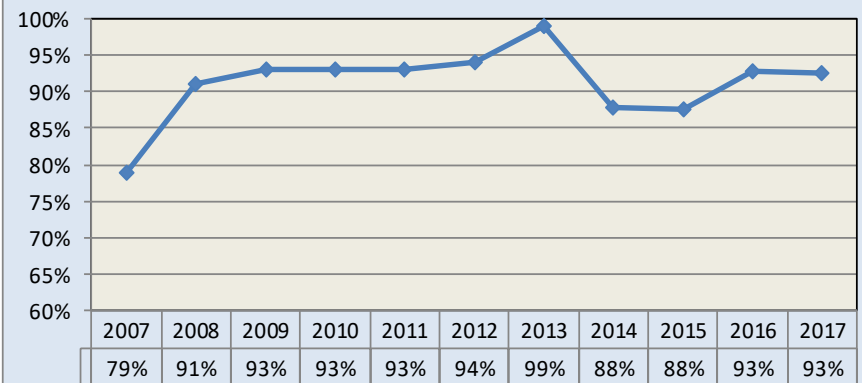
User Survey Results: Allocation Process

- Positive scores overall. Many thoughtful comments were provided through free-form text submissions. Thank you for your input.

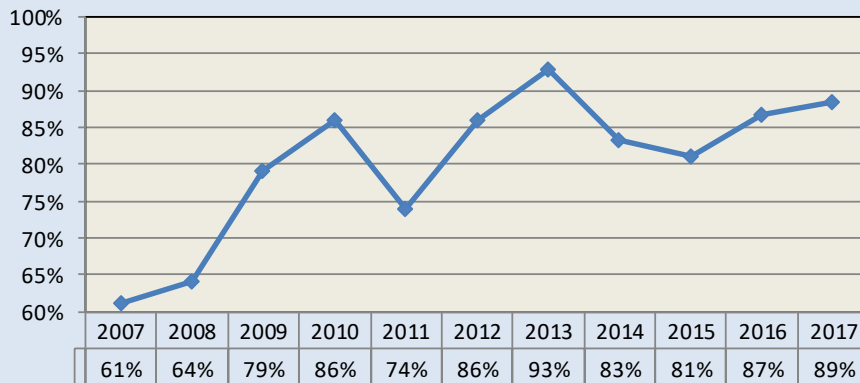
Overall Satisfaction with the Allocation Process



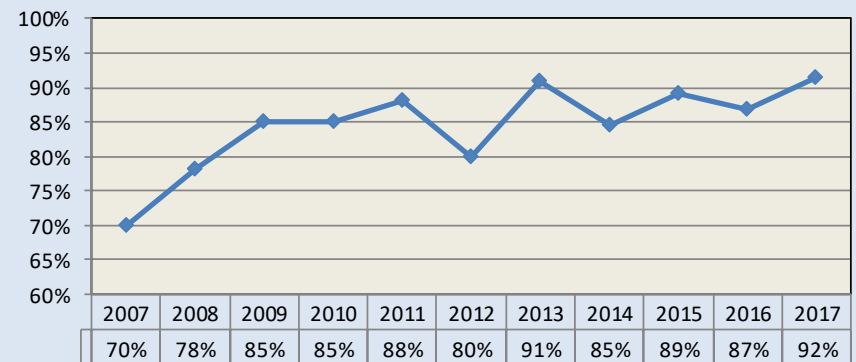
Clarity of the Call for Proposals



Transparency of the Allocation Process

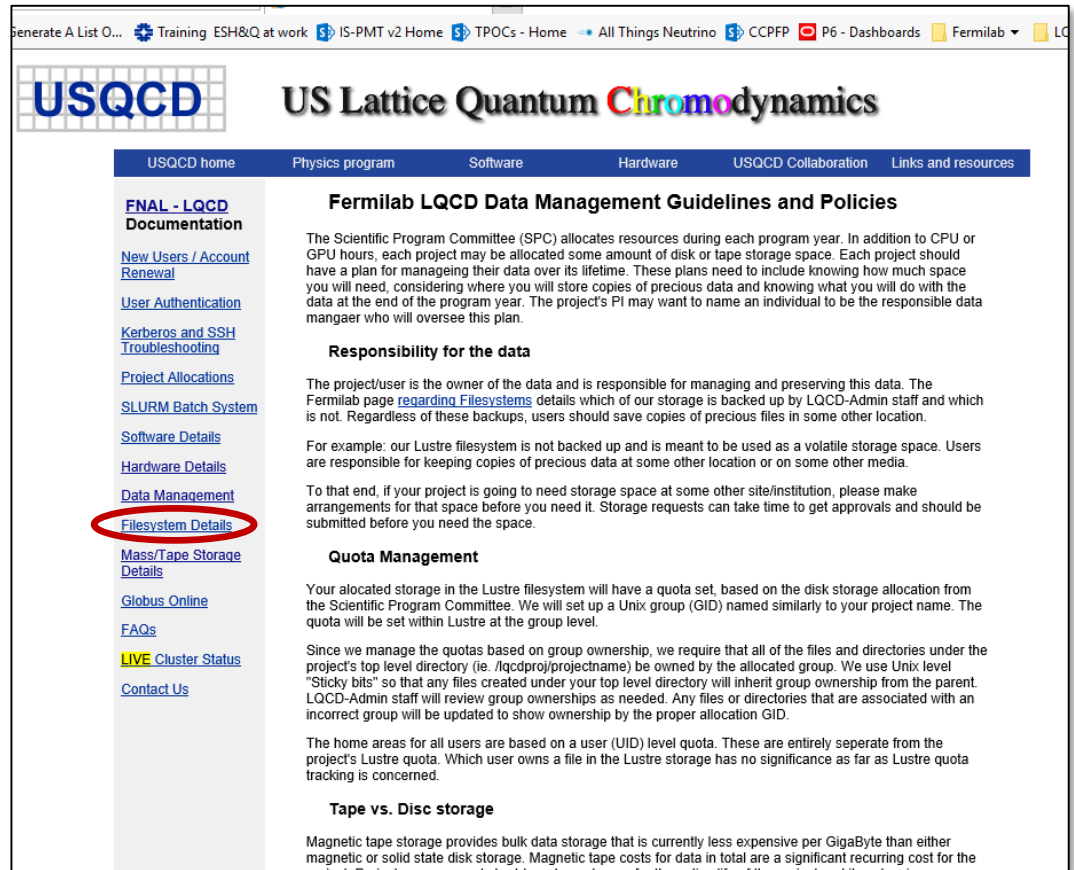


Allocation Process Helps Maximize Scientific Output



Data Management Policies and Guidelines

- ▶ Active data management is becoming more important as data storage needs (disk and tape) continue to grow and budgets remain fixed.
- ▶ Site managers for the three host labs have been creating or reviewing data management policies and guidelines for short-term data storage.
 - JLab data management policy last updated Jul 1, 2014
 - FNAL data management policy updated Apr 16, 2019
 - BNL data management policy under development



The screenshot shows the USQCD website with the following content:

- USQCD logo and US Lattice Quantum Chromodynamics title.
- Navigation menu: USQCD home, Physics program, Software, Hardware, USQCD Collaboration, Links and resources.
- Left sidebar menu:
 - FNAL - LQCD Documentation
 - New Users / Account Renewal
 - User Authentication
 - Kerberos and SSH Troubleshooting
 - Project Allocations
 - SLURM Batch System
 - Software Details
 - Hardware Details
 - Data Management
 - Filesystem Details** (circled in red)
 - Mass/Tape Storage Details
 - Globus Online
 - FAQs
 - LIVE Cluster Status
 - Contact Us
- Main content area:
 - Fermilab LQCD Data Management Guidelines and Policies**
 - The Scientific Program Committee (SPC) allocates resources during each program year. In addition to CPU or GPU hours, each project may be allocated some amount of disk or tape storage space. Each project should have a plan for managing their data over its lifetime. These plans need to include knowing how much space you will need, considering where you will store copies of precious data and knowing what you will do with the data at the end of the program year. The project's PI may want to name an individual to be the responsible data manager who will oversee this plan.
 - Responsibility for the data**
 - The project/user is the owner of the data and is responsible for managing and preserving this data. The Fermilab page [regarding Filesystems](#) details which of our storage is backed up by LQCD-Admin staff and which is not. Regardless of these backups, users should save copies of precious files in some other location.
 - For example: our Lustre filesystem is not backed up and is meant to be used as a volatile storage space. Users are responsible for keeping copies of precious data at some other location or on some other media.
 - To that end, if your project is going to need storage space at some other site/institution, please make arrangements for that space before you need it. Storage requests can take time to get approvals and should be submitted before you need the space.
 - Quota Management**
 - Your allocated storage in the Lustre filesystem will have a quota set, based on the disk storage allocation from the Scientific Program Committee. We will set up a Unix group (GID) named similarly to your project name. The quota will be set within Lustre at the group level.
 - Since we manage the quotas based on group ownership, we require that all of the files and directories under the project's top level directory (ie. /lqcdproj/projectname) be owned by the allocated group. We use Unix level "Sticky bits" so that any files created under your top level directory will inherit group ownership from the parent. LQCD-Admin staff will review group ownerships as needed. Any files or directories that are associated with an incorrect group will be updated to show ownership by the proper allocation GID.
 - The home areas for all users are based on a user (UID) level quota. These are entirely separate from the project's Lustre quota. Which user owns a file in the Lustre storage has no significance as far as Lustre quota tracking is concerned.
 - Tape vs. Disc storage**
 - Magnetic tape storage provides bulk data storage that is currently less expensive per GigaByte than either magnetic or solid state disk storage. Magnetic tape costs for data in total are a significant recurring cost for the project. Project managers should take time to plan for the entire life of the project and its data.

Recent success story. Working with numerous PIs, Ken Schumacher was able to facilitate the recovery of 330 TB of disk space that was being used to store legacy data. All of the legacy data was moved or deleted as appropriate. Recovering ~40% of our available storage allocation allows us to operate more efficiently and better satisfy current needs.

Upcoming DOE Science Review

- ▶ No annual project review this year.
- ▶ In July, DOE/HEP will be reviewing a proposal to extend the LQCD Research Program to the next 5-year period (Oct 2019 through Sep 2024).
 - What is the scientific case for continuing simulations of QCD in high energy physics past 2019? Are the goals of the proposed research program aligned with the experimental and theoretical physics goals of HEP for the period 2020–2024.
 - What is the impact and interplay of lattice QCD simulations on the experimental and theoretical programs of HEP? Will the value of our experimental and theoretical program be measurably enhanced by such simulations?
 - Why is an extended project needed if ASCR* is providing the lattice community access to Leadership Class machines? In particular, is mid-scale hardware, such as CPU and GPU Institutional Clusters, essential and cost effective in such an environment? What is the optimal mix of machines, Leadership Class and mid-scale clusters, given realistic budget scenarios?
 - What are the plans for Fermilab and Brookhaven for LQCD Institutional Cluster computing? How are these plans incorporated into your proposal for LQCD research programs in 2020–2024?
- ▶ Review panel comprises computational scientists and high energy theoretical and experimental physicists.
- ▶ Review planning and preparations are actively underway.
- ▶ Brookhaven and Fermilab are both committed to supporting and expanding mid-scale Institutional Cluster hardware and related storage systems.

*DOE Office of Advanced Scientific Computing Research

***Thank you for keeping
our systems busy!***

Questions?