# Fermilab Facilities report

Amitoj Singh

USQCD All-Hands Collaboration Meeting

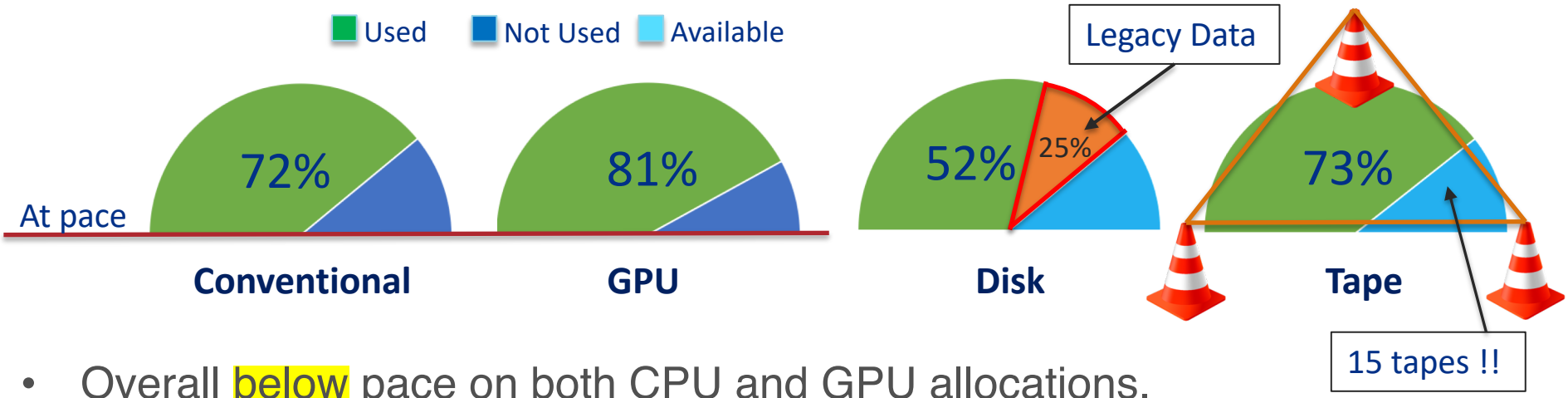
20-21 April 2018

# Hardware – Current Clusters

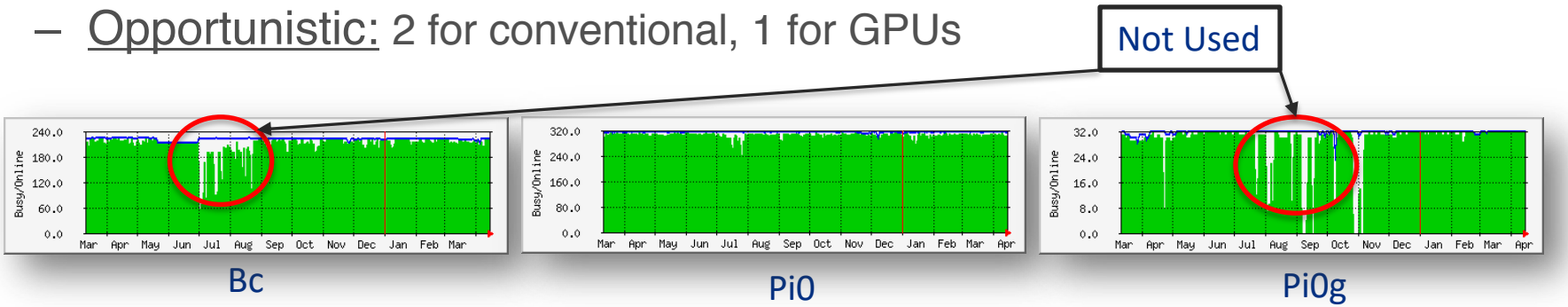
| Name | CPU | Nodes | Cores GPUs | Network | Equivalent Jpsi core or Fermi gpu-hrs | Online |
|---|---|---|---|---|---|---|
| Ds* | Quad 2.0 GHz Opteron 6128 (8-core) | 179 | 5,728 | Infiniband QDR | Unallocated | Dec 2010 Aug 2011 |
| *Dsg** | *Dual NVIDIA M2050 GPUs+Intel 2.53 GHz E5630 (4-core)* | *39* | *312 Cores ?? GPUs* | *Infiniband QDR* | *Unallocated* | *Mar 2012* |
| Bc | Quad 2.8 GHz Opteron 6320 (8-core) | 222 | 7,104 | Infiniband QDR | **1.48** Jpsi | July 2013 |
| Pi0 | Dual 2.6 GHz Xeon E2650v2 (8-core) | 314 | 5,024 | Infiniband QDR | **3.14** Jpsi | Oct 2014 Apr 2015 |
| *Pi0g* | *Dual NVIDIA K40 GPUs+Intel 2.6 GHz E2650v2 (8-core)* | *32* | *512 Cores 128 GPUs* | *Infiniband QDR* | ***2.6** Fermi* | *Oct 2014* |
| | ***TOTAL*** | 786 | 18,680 Cores 128 GPUs | | | |

* Unallocated resource

# Progress Against Allocations as of 4-16-2018




- Overall below pace on both CPU and GPU allocations.
- PY17-18 Allocation status:
  - Class A (17 total): 2 finished, 6 at or above pace
  - Class B (2 total): 1 at or above pace
  - Class C: 3 for conventional
  - Opportunistic: 2 for conventional, 1 for GPUs






Fermilab

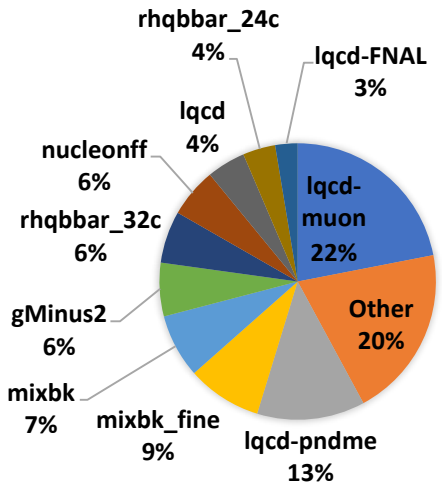
# Storage

- Global disk storage:
  - 776 TB Lustre file-system at /lqcdproj. (not backed up)
  - 197 TB Lustre file-system at /lfsz. (not backed up)
  - 14.5 TB “project” space at /project (backed up nightly)
  - 6 GB per user at /home on each cluster (backed up nightly)
- Robotic tape storage is available via *dccp* commands against the dCache filesystem at /pnfs/lqcd.
  - Some users will benefit from direct access to tape by using *encp* commands on lqcdsrm.fnal.gov
  - Please email us if you plan on transferring a large number of small (< 200MB) or big (>1TB) files. We may want to guide how these are written to avoid wasted space on tape.
- Globus Online endpoint:
  - lqcd#fnal - for transfers in or out of our Lustre file system.
- Data integrity is your responsibility.

# Tape Storage

Urgent push to migrate data on existing tape media to LTO8 !! Estimated cost of migrating ALL data is ~$69k. Please review and delete legacy data under /pnfs/lqcd.

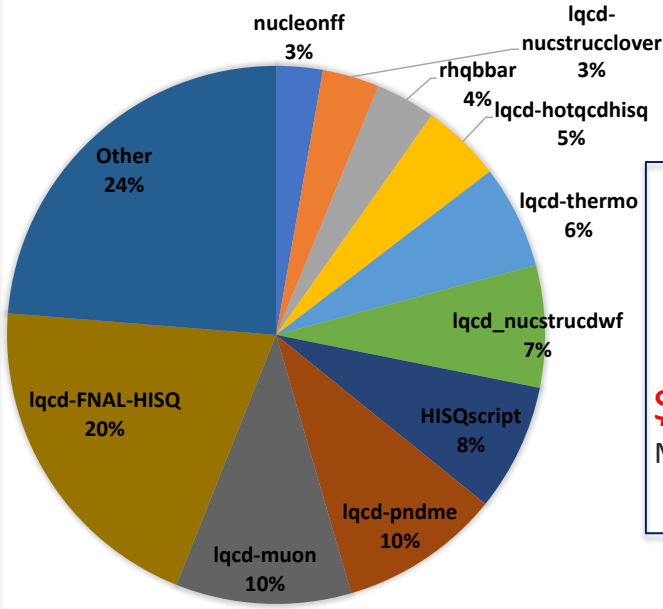
LTO4 Media (0.8TB)
Total: ~1.3 PB, 1,617 tapes




2004-2012

Maintenance
$5.6k/year
Migration to LTO8
$6,400

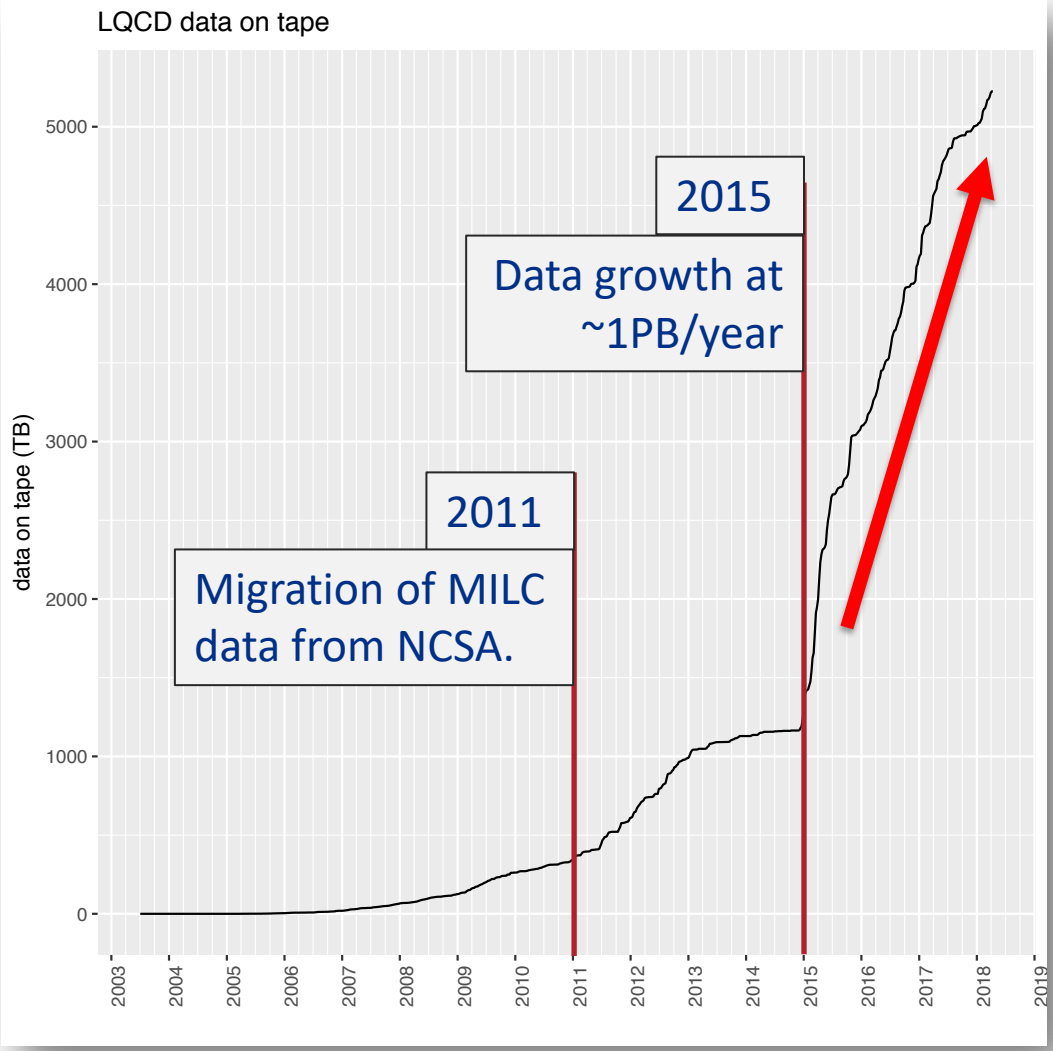
T10K Media (8.4TB),
Total: ~4.2 PB, 564 tapes




2012-YTD

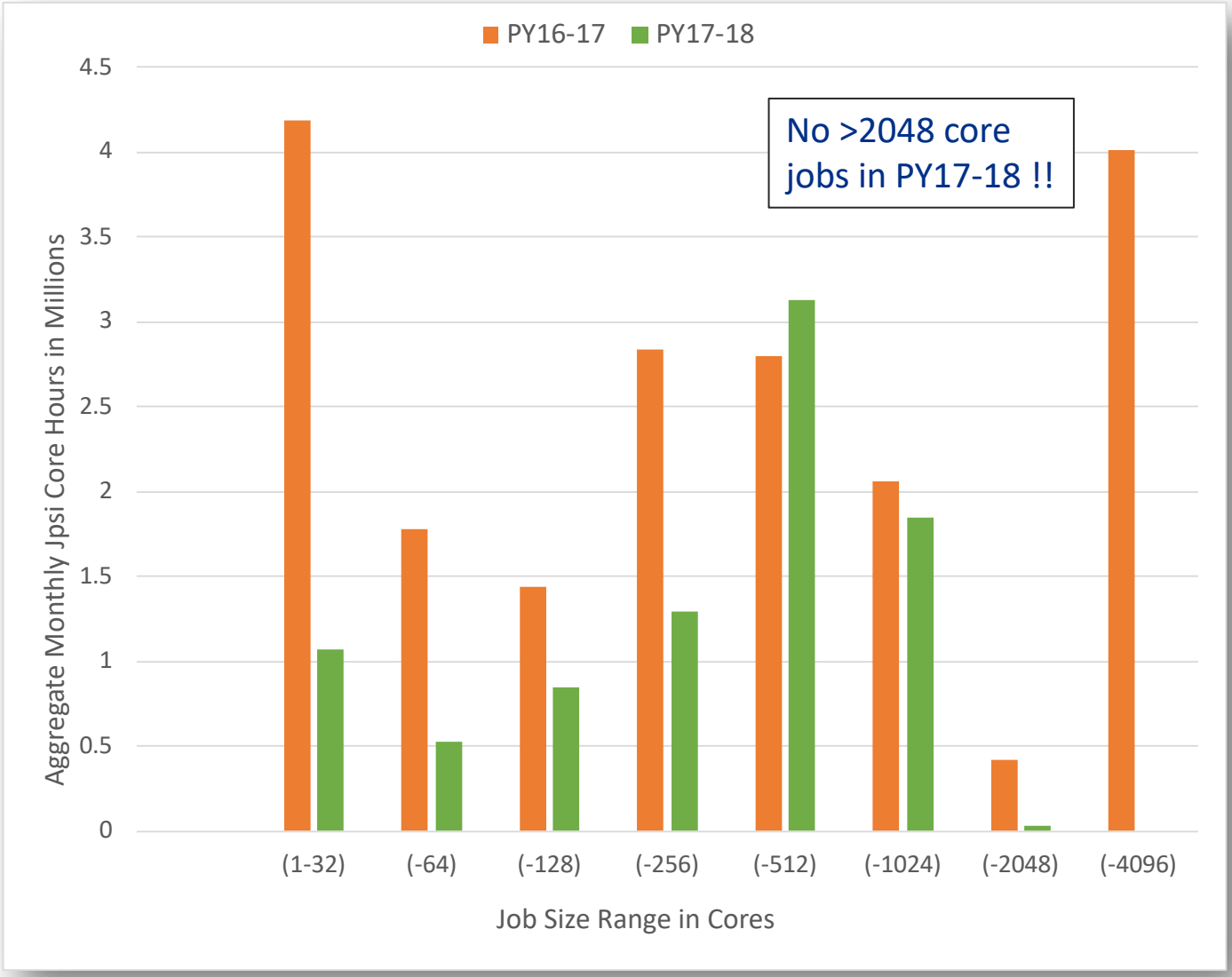
Maintenance
$12.4k/year
Migration to LTO8
$63,000

# Tape Storage




The rate at which data is accumulating beyond 2015, with no data management plan, is worrisome !!

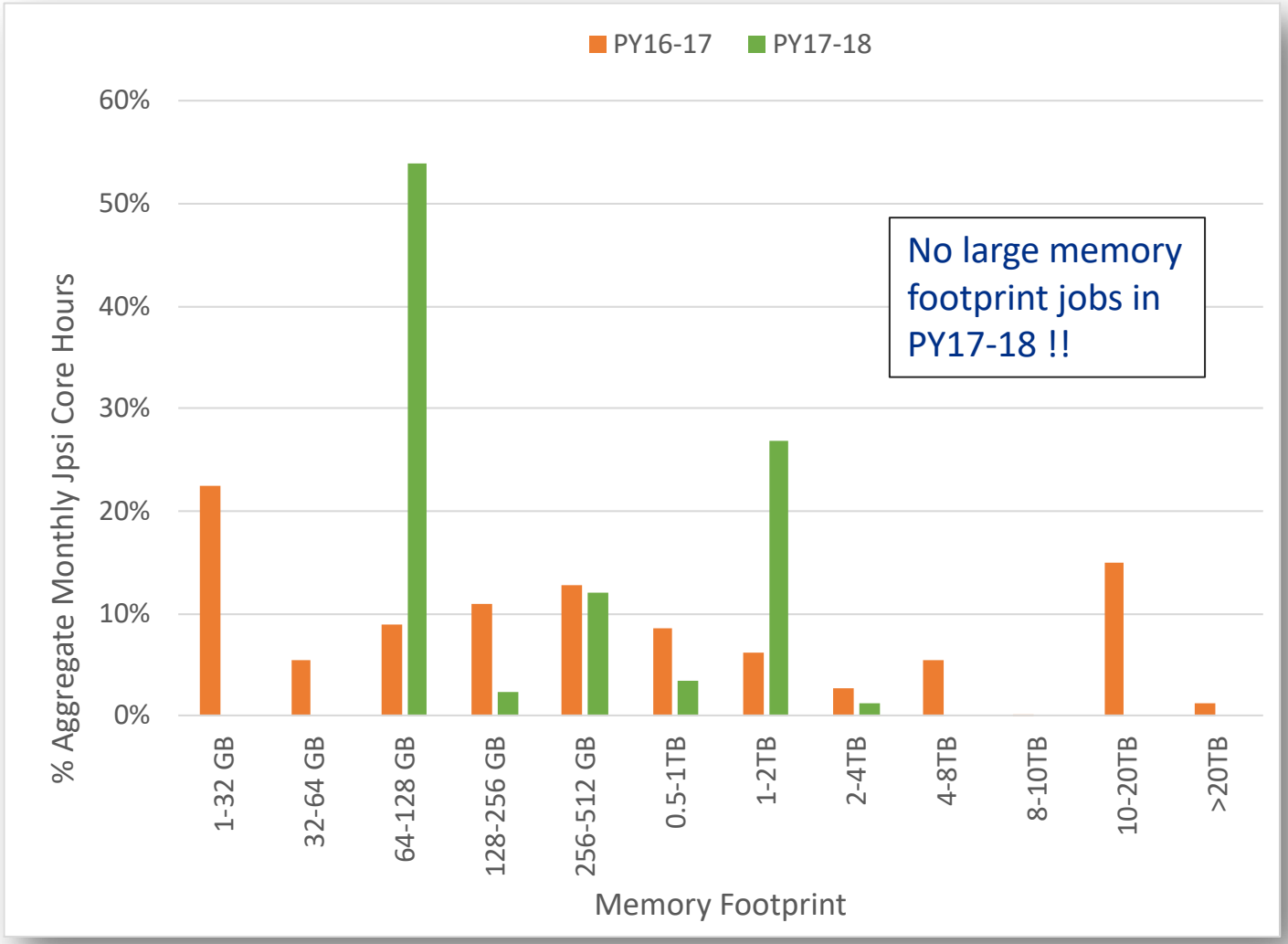
# Changes in Job Size Distribution






Fermilab

# Changes in Job Memory Footprint






Fermilab

# New web portal for USQCD Fermilab resources

https://www.usqcd.org/fnal/clusterinfo.html

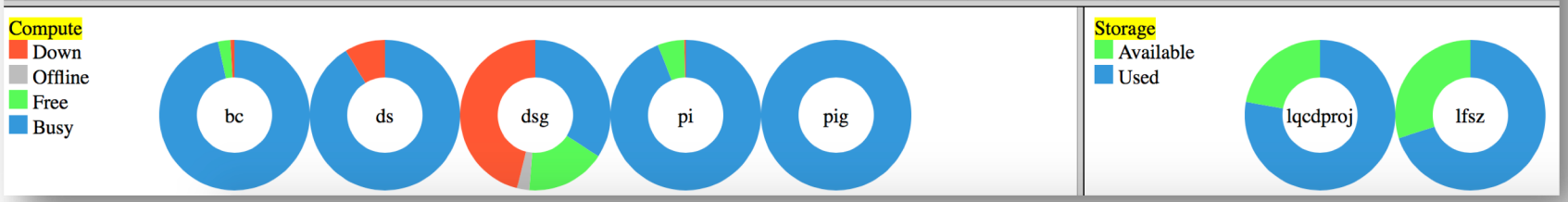
pi0s.fnal.gov | bc2.fnal.gov | ds2.fnal.gov

Refresh Rate=5 mins. Last updated on: 04-17-2018 4:25 PM

**PBS job status for pios.fnal.gov**

| Job ID | User | Submitted On | Run On | Time Used | Time Req | Status ▲ | Queue | Nodes | Project |
|---|---|---|---|---|---|---|---|---|---|
| 810987 | Anna Hasenfratz | 04-17-18 4:08 PM | 04-17-18 4:17 PM | 00:13:55 | 03:00:00 | R | pi | 2 | fourpluseight |
| 810986 | Anna Hasenfratz | 04-17-18 3:35 PM | 04-17-18 3:42 PM | 00:46:56 | 03:00:00 | R | pi | 2 | fourpluseight |
| 810983 | Raghav Jha | 04-17-18 12:50 PM | 04-17-18 2:14 PM | 02:07:46 | 24:00:00 | R | pi | 1 | latticesusy |
| 810982 | Raghav Jha | 04-17-18 12:50 PM | 04-17-18 1:02 PM | 03:31:25 | 24:00:00 | R | pi | 1 | latticesusy |
| 810966 | Yuzhi Liu | 04-17-18 8:19 AM | 04-17-18 3:22 PM | 01:08:19 | 24:00:00 | R | pi | 16 | qcd-qed |
| 810965 | Yuzhi Liu | 04-17-18 8:19 AM | 04-17-18 2:27 PM | 02:02:49 | 24:00:00 | R | pi | 16 | qcd-qed |
| 810940 | Nouman Butt | 04-17-18 3:17 AM | 04-17-18 1:02 PM | 05:39:59 | 24:00:00 | R | pi | 4 | latticesusy |
| 810939 | Nouman Butt | 04-17-18 3:17 AM | 04-17-18 1:02 PM | 05:39:59 | 24:00:00 | R | pi | 4 | latticesusy |
| 810938 | Nouman Butt | 04-17-18 3:17 AM | 04-17-18 1:02 PM | 05:39:58 | 24:00:00 | R | pi | 4 | latticesusy |
| 810937 | Nouman Butt | 04-17-18 3:17 AM | 04-17-18 1:02 PM | 05:39:58 | 24:00:00 | R | pi | 4 | latticesusy |
| 810879 | Yuzhi Liu | 04-16-18 10:18 PM | 04-17-18 1:02 PM | 17:17:05 | 24:00:00 | R | pi | 16 | qcd-qed |
| 810878 | Raghav Jha | 04-16-18 8:01 PM | 04-17-18 1:02 PM | 17:17:06 | 24:00:00 | R | pi | 2 | latticesusy |
| 810866 | Raghav Jha | 04-16-18 4:04 PM | 04-17-18 1:02 PM | 17:58:52 | 24:00:00 | R | pi | 3 | latticesusy |
| 810865 | Raghav Jha | 04-16-18 4:04 PM | 04-17-18 1:02 PM | 17:58:54 | 24:00:00 | R | pi | 3 | latticesusy |
| 810864 | Raghav Jha | 04-16-18 4:04 PM | 04-17-18 1:02 PM | 17:58:49 | 24:00:00 | R | pi | 6 | latticesusy |
| 810863 | Raghav Jha | 04-16-18 4:04 PM | 04-17-18 1:02 PM | 18:17:11 | 24:00:00 | R | pi | 2 | latticesusy |
| 810862 | Raghav Jha | 04-16-18 4:04 PM | 04-17-18 1:02 PM | 17:58:52 | 24:00:00 | R | pi | 3 | latticesusy |
| 810847 | David Schaich | 04-16-18 1:01 PM | 04-17-18 2:10 PM | 02:14:29 | 20:00:00 | R | pi | 4 | latticesusy |

Links to usage reports coming soon.

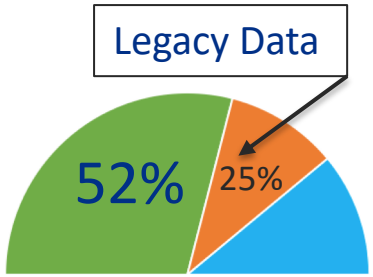





Fermilab

# Upcoming upgrades and major changes

- Bc cluster:
  - For the 2018-19 program year, the Bc cluster will be available to you as an unallocated resource along with Ds and Dsg. Drop in Jpsi-equivalent core-hours from 190M to 113.6M. Available GPU hours remain the same i.e. 2.4M GPU-hours.
- Batch queueing system
  - At the start of the new allocation year (July 1, 2018), we will be replacing Torque/Maui with SLURM. Plan is to have all three USQCD sites use SLURM.
- Data Preservation Policy:
  - Legacy disk data not covered by a storage allocation and not community owned should, within 30 days from the end of the projects' allocation, either be moved off site or to tape by project. If no action is taken within the 30 days, data will be archived at the site's discretion unless prior arrangements have been made.



# User Support

Fermilab points of contact:

**Please avoid sending support related emails directly to the POCs.**

- Rick Van Conant, vanconant@fnal.gov
- Alex Kulyavtsev, aik@fnal.gov (Mass Storage and Lustre)
- Ken Schumacher, kschu@fnal.gov
- Jim Simone, simone@fnal.gov
- Amitoj Singh, amitoj@fnal.gov
- Alexei Strelchenko, astrel@fnal.gov (GPUs)

**Please use lqcd-admin@fnal.gov for incidents only as this autogenerates a support ticket.**

**Please use hpc-admin@fnal.gov for requests or general correspondence.**

# Questions?