Fermilab Facilities report

Gerard Bernabeu Altayo

USQCD All-Hands Collaboration Meeting

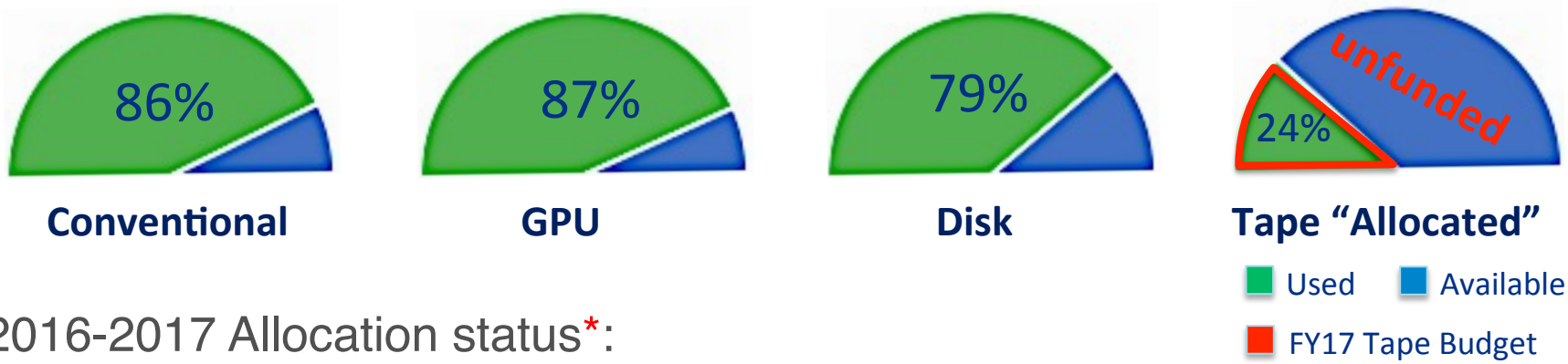
28-29 April 2017

Hardware – Current Clusters

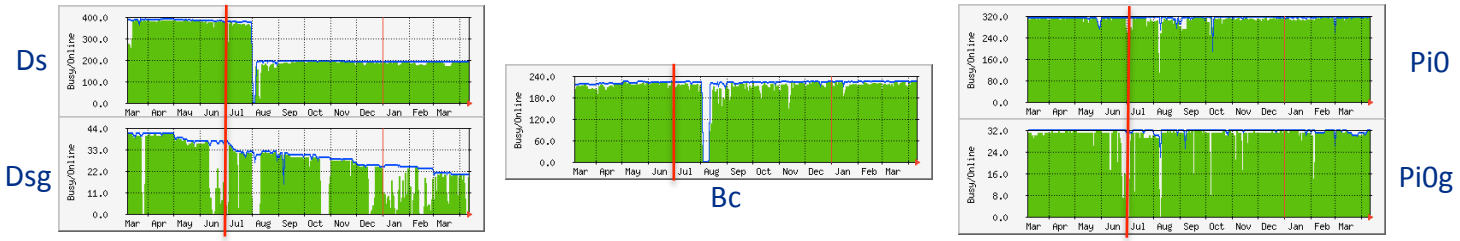
Name	CPU	Nodes	Cores GPUs	Network	Equivalent Jpsi core or Fermi gpu-hrs	Online
Ds*	Quad 2.0 GHz Opteron 6128 (8-core)	196	6,272	Infiniband QDR	**1.33** Jpsi	Dec 2010 Aug 2011
*Dsg**	*Dual NVIDIA M2050 GPUs+Intel 2.53 GHz E5630 (4-core)*	*20*	*160 Cores 40 GPUs*	*Infiniband QDR*	***1.1** Fermi*	*Mar 2012*
Bc	Quad 2.8 GHz Opteron 6320 (8-core)	224	7,168	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	19,136 Cores 168 GPUs			

* Unallocated resource

Progress Against Allocations



- 2016-2017 Allocation status*:
  - Class A (21 total): 3 finished, 7 at or above pace
  - Class B (3 total): 1 at or above pace
  - Class C: 3 for conventional
  - Opportunistic: 4 for conventional, 3 for GPUs



* as of 4/13/2017

Fermilab

Storage

- Global disk storage:
  - 782 TB Lustre file-system at /lqcdproj.
  - 197 TB Lustre file-system at /lfsz.
  - 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.
  - Please email us if writing TB-sized files. With 8.5TB tapes, we may want to guide how these are written to avoid wasted space.
  - Remote direct access to dCache is available via GridFTP (no Globus Online support)
- Worker nodes have local storage at /scratch.
- Globus Online endpoint:
  - lqcd#fnal - for transfers in or out of our Lustre file system.

Storage – Data integrity

- Some friendly reminders:
  - **Data integrity is your responsibility.**
  - With the exception of /home area and /project, backups are not performed.
  - Make copies on different storage hardware of any of your critical data.
  - Data can be copied to tape using *dccp* or *encp* commands. Please contact us for details. We have never lost LQCD data on Fermilab tape.
  - At 45 disk pools and growing on Lustre, the odds of a partial failure will eventually catch up with us.

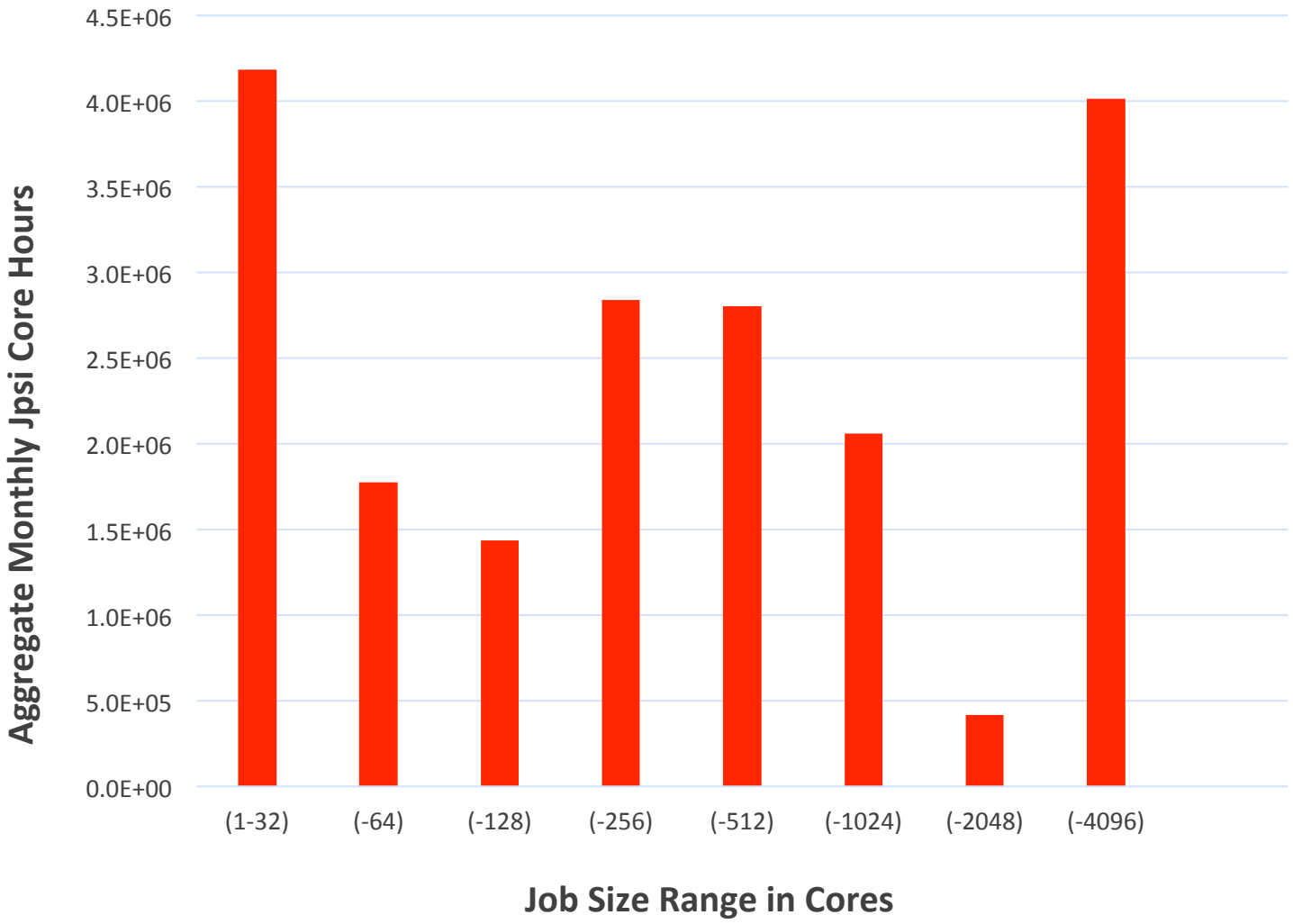
Lustre File-System

- Lustre Statistics:
  - Capacity: 979 TB available, 777 TB used (79% used)
  - Files: 126 million (76M last year)
  - File sizes: largest file is 489 GB (tar ball) 230 GB (file), average size is 6.7 MB

- Please email us if writing TB-sized files. For Lustre there will be tremendous benefit in striping such files across several OSTs both for performance and for balancing space used per storage target.

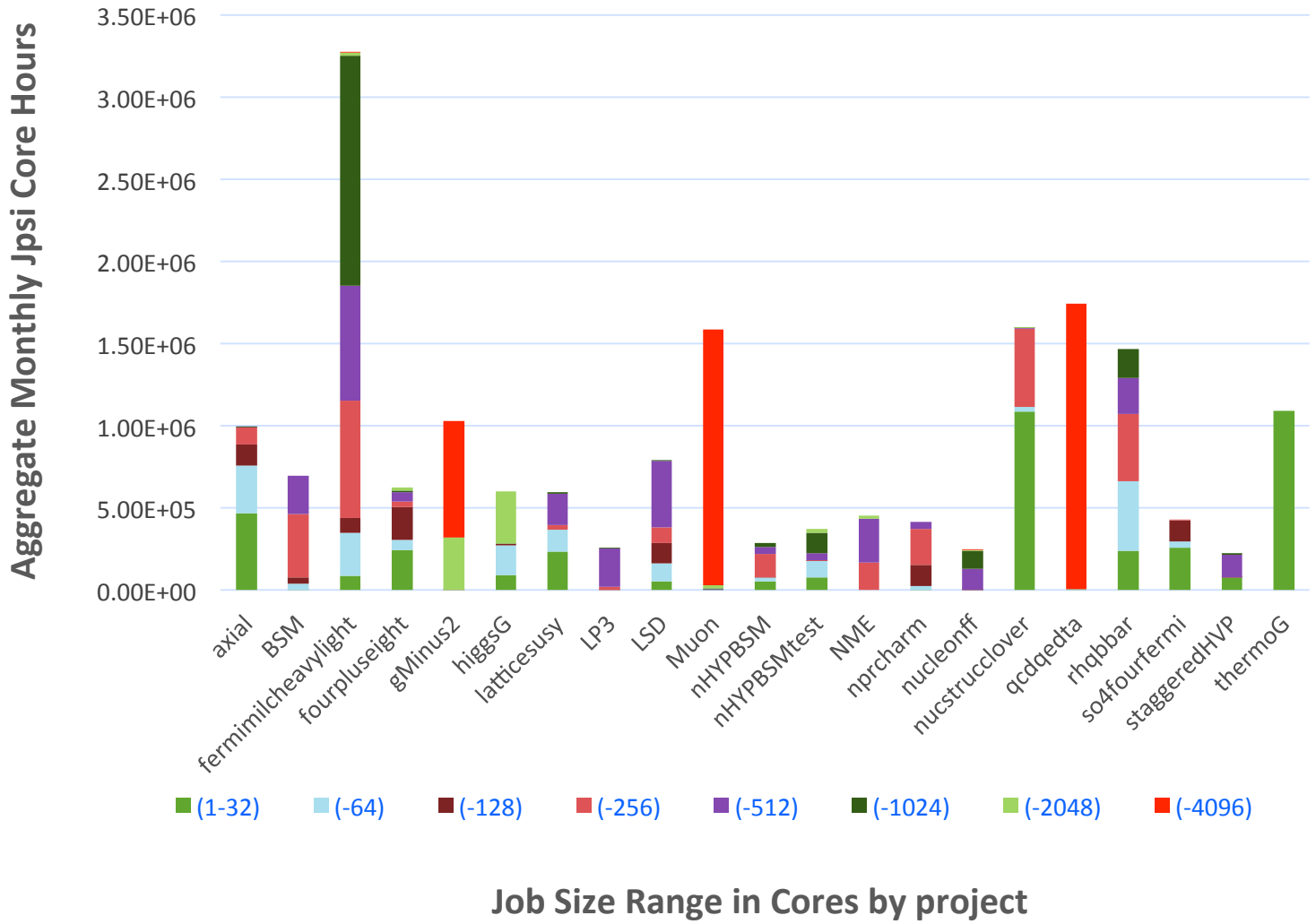
- NOTE: No budget till FY18 to grow disk storage capacity. Please remove or migrate old data off FNAL disk storage.

USQCD 2016-17 Fermilab Clusters Job Size Statistics



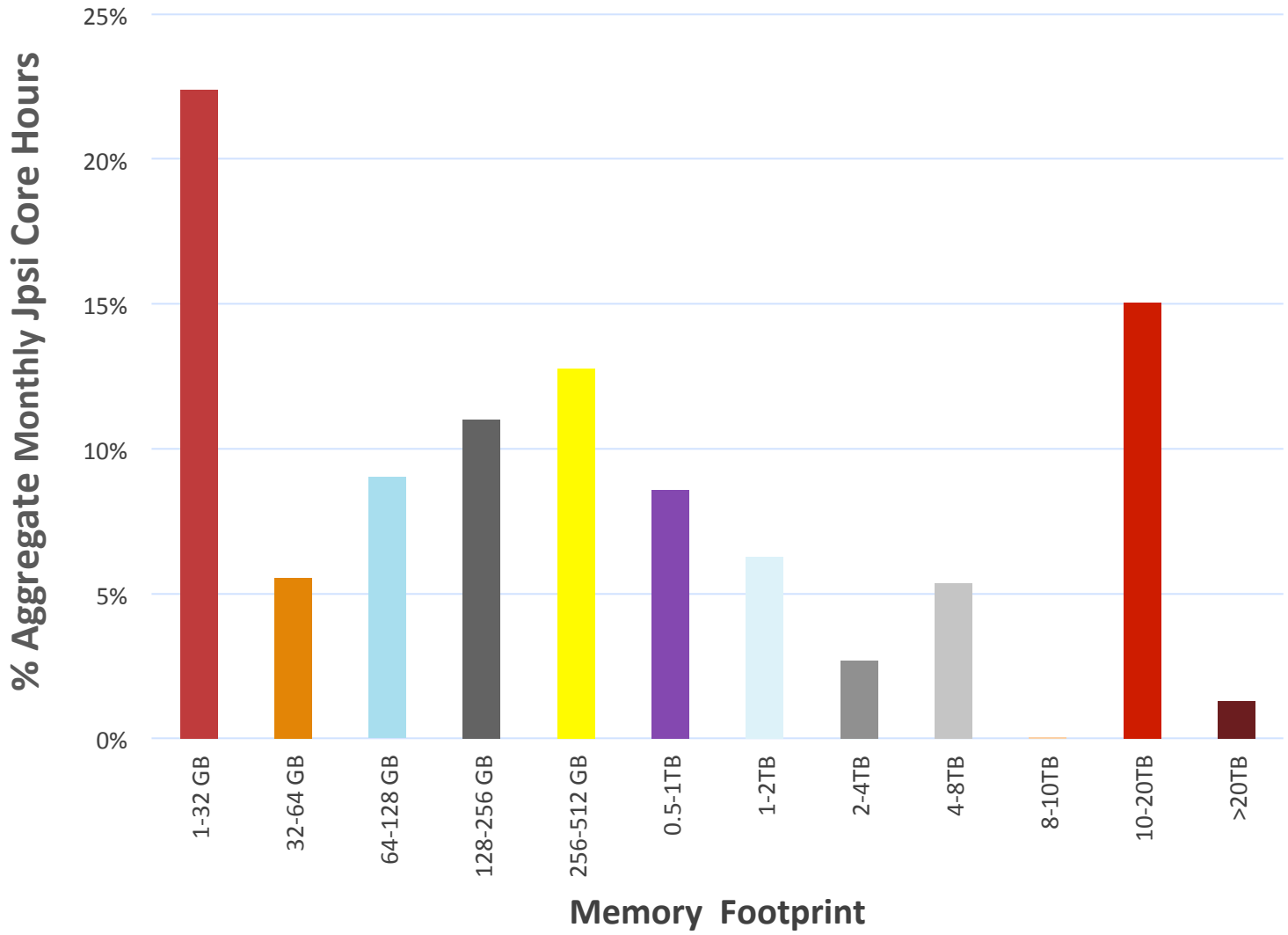
Fermilab

USQCD 2016-17 Fermilab Cluster Job Size Statistics



Fermilab

USQCD 2016-17 Fermilab Clusters Job Memory Footprint Statistics



Fermilab

Upcoming upgrades and major changes

- Ds and Dsg clusters:
  - For the 2017-18 program year, the Ds and Dsg clusters will be available to you as an unallocated resource.
  - As of now there are 196 Ds and 20 Dsg worker nodes in good to fair condition. There is a tentative plan to reconfigure Dsg worker nodes with failed GPUs as conventional worker nodes.
- Data Preservation Policy:
  - 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. 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 use lqcd-admin@fnal.gov for incidents or requests.

Please avoid sending support related emails directly to the POCs.

- Gerard Bernabeu, gerard1@fnal.gov
- Rick Van Conant, vanconant@fnal.gov
- Alex Kulyavtsev, aik@fnal.gov (Mass Storage and Lustre)
- Paul Mackenzie, mackenzie@fnal.gov
- Ken Schumacher, kschu@fnal.gov
- Jim Simone, simone@fnal.gov
- Amitoj Singh, amitoj@fnal.gov
- Alexei Strelchenko, astrel@fnal.gov (GPUs)

Questions?