Report on the Clusters at Fermilab

Don Holmgren

USQCD All-Hands Meeting

BNL

April 19-20, 2013

Outline

- Hardware
- Storage
- Statistics
- Summer Loadsheds

Hardware – Current and Next Clusters

Name	CPU	Nodes	Cores	Network	DWF	Asqtad	Online
Jpsi (2008) (2009)	Dual 2.1 GHz Opteron 2352 (Quad Core)	856	6848	Infiniband Double Data Rate	10061 MFlops per Node	9563 MFlops per Node	Jan 2009 / Apr 2009 8.40 TFlops
Ds (2010) (2011)	Quad 2.0 GHz Opteron 6128 (8 Core)	421	13472	Infiniband Quad Data Rate	51.2 GFlops per Node	50.5 GFlops per Node	Dec 2010 11 TFlops Aug 2011 21.5 TF
Dsg (2012)	NVIDIA M2050 GPUs + Intel 2.53 GHz E5630 (quad core)	76	152 GPUs 608 Intel	Infiniband Quad Data Rate	29.0 GFlops per Node (cpu)	17.2 GFlops per Node (cpu)	Mar 2012
Bc (2013)	Quad 2.8 GHz Opteron 6320 (8 Core)	224	7168	Infiniband Quad Data Rate	*53.3 Gflops per Node (est.)*	*55.9 Gflops per Node (est.)*	*June 2013*

New Cluster (Bc)

- Hardware design:
  - Similar to Ds
  - Nodes are quad-socket 8-core AMD 6320 (*Abu Dhabi*), 2.8 GHz
  - 2 GB/core → 64 GB/node
  - QDR Infiniband, full bandwidth
  - 224 servers
  - Estimated 12.6 TF sustained (DWF:asqtad inverter average using jobs with 128 cores)
- Delivery scheduled for early May
  - Friendly user testing could start by May 15
  - Release to production as early as June 15

Storage

- Global disk storage:
  - 614 TiB Lustre filesystem at /lqcdproj
  - ~ 6 TiB “project” space at /project (backed up nightly)
  - ~ 6 GiB 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 using *encp* on lqcdsrm.fnal.gov
- Worker nodes have local storage at /scratch
  - Multi-node jobs can specify combining /scratch from one or more nodes into /pvfs
  - /pvfs is visible to all nodes of the job and is deleted at job end

Storage

- We have setup two Globus Online endpoints:
  - usqcd#fnal – for transfers directly into our out of FNAL's robotic tape system. Use DOE or OSG certificates, or Fermilab KCA certificates. You must become a member of either the FNAL LQCD VO or the ILDG VO.
  - lqcd#fnal – for transfers into our out of our Lustre file system (/lqcdproj). You must use a FNAL KCA certificate. See http://www.usqcd.org/fnal/globusonline.html
  - You may also use GlobusConnect to setup a personal endpoint
- We have provisioned two machines with 10 gigE connections
  - lqcdgo.fnal.gov – used for Globus Online, not available for interactive use
  - lqcdsrm.fnal.gov – best machine to use for moving data to/from tape.

lqcdsrm.fnal.gov

- We have provisioned this node with a 10gigE connection to the FNAL WAN. It also has an Infiniband QDR connection to our Lustre filesystem, and access to the home areas of all of the clusters.
- *encp* commands interact directly with tape drives and with the "Small Files Aggregation" facility. All new tape drives are Oracle T10K, with 240 MB/sec transfer rates.
  - if your files are "small" (100 MB or less), please contact us about configuring your project to use the SFA
- *dccp* commands interact with disk cache nodes with only 1gigE interfaces, so you must have multiple commands in flight to take advantage of the 10gigE. For example,

```
cat filelist | parallel -j 10 dccp {} /pnfs/lqcd/…
```

Storage – Lustre Statistics

- 614 TiB capacity, 540 TiB currently used, 114 disk pools
  (2012: 543 TiB capacity, 475 TiB used in 114 pools)
- 101M files (101M last year)
  However, 42.7M files that were in /lqcdproj/project have been moved to an NFS export
- File sizes: 344.8 GiB maximum (a log file!)
  5.73 MiB average (8.54 MiB last year)
- Directories:323K (321K last year excluding /project)
  799K files in largest directory!

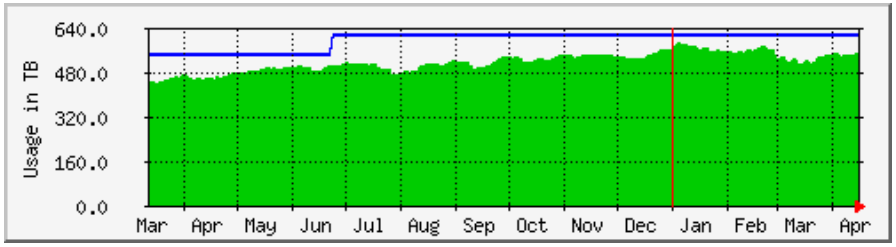
Storage – Planned Changes

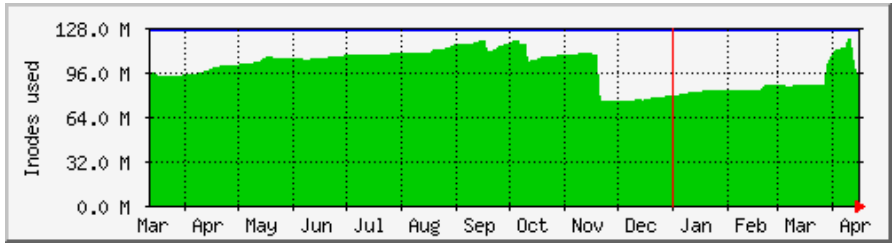
1. Deploy additional Lustre storage – now and during 2013 project year
   - Expect to add about ~ 116 TiB by late-May (total to 730 TB)
   - Will hold another ~ 116 TiB in reserve for addition after the new project year starts
   - Once funds are available in FY14, expect to add another ~ 230 TiB (however, continuing budget resolutions always delay this type of spending)
2. Will schedule a downtime near the beginning of the new project year to expand metadata space (this controls file and directory capacity)

Storage – Date Integrity

- Some reminders:
  - Data integrity is your responsibility
  - With the exception of home areas and /project, backups are not performed
  - Make copies on different storage hardware of any of your data that are critical
  - Data can be copied to tape using dccp commands. Please contact us for details. We can also show you how to make multiple copies that are guaranteed to be on different tapes. We have never lost LQCD data on Fermilab tape (1.60 PiB and growing).
  - At 114 disk pools and growing, the odds of a partial failure will eventually catch up with us

Storage - Utilization

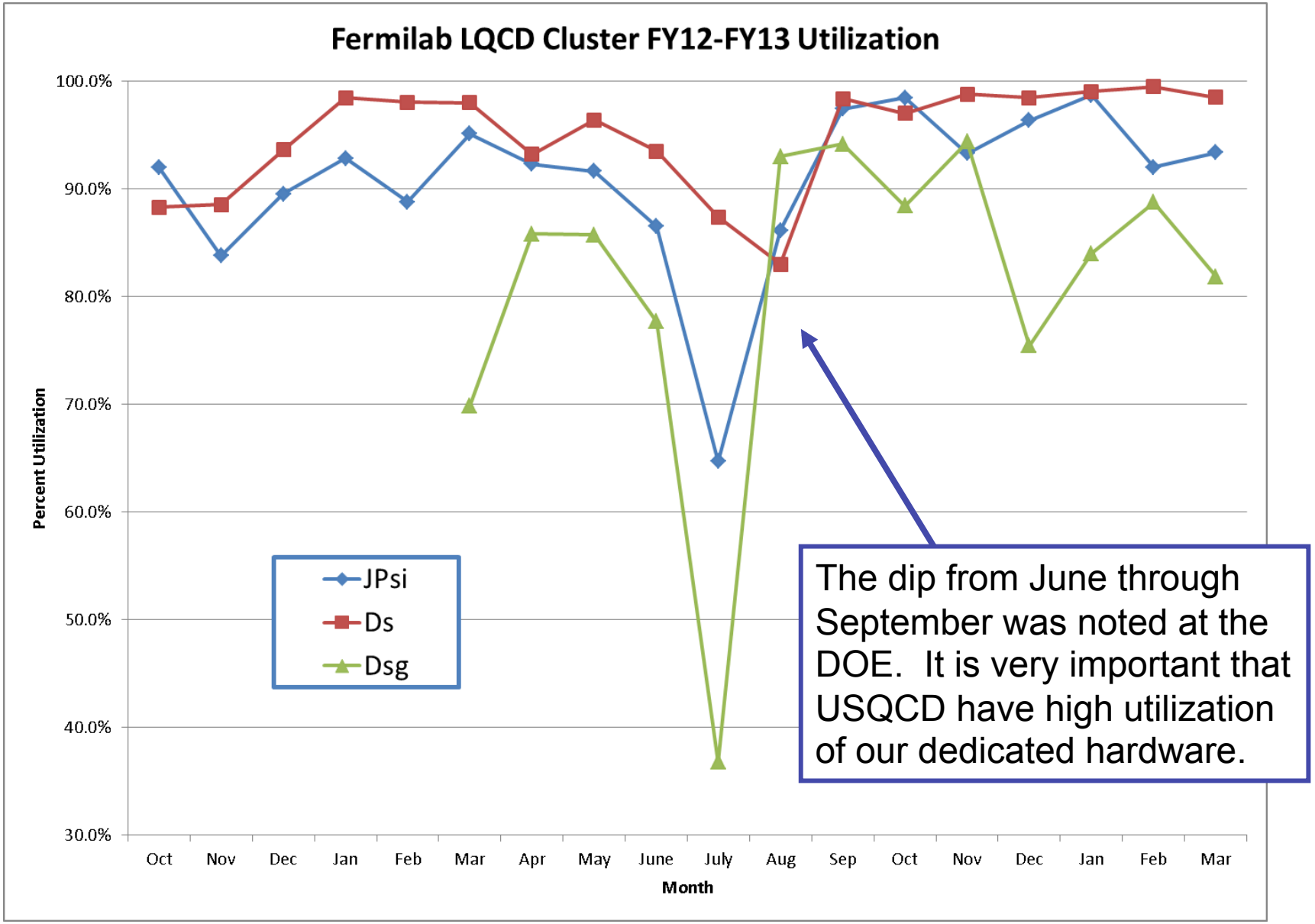


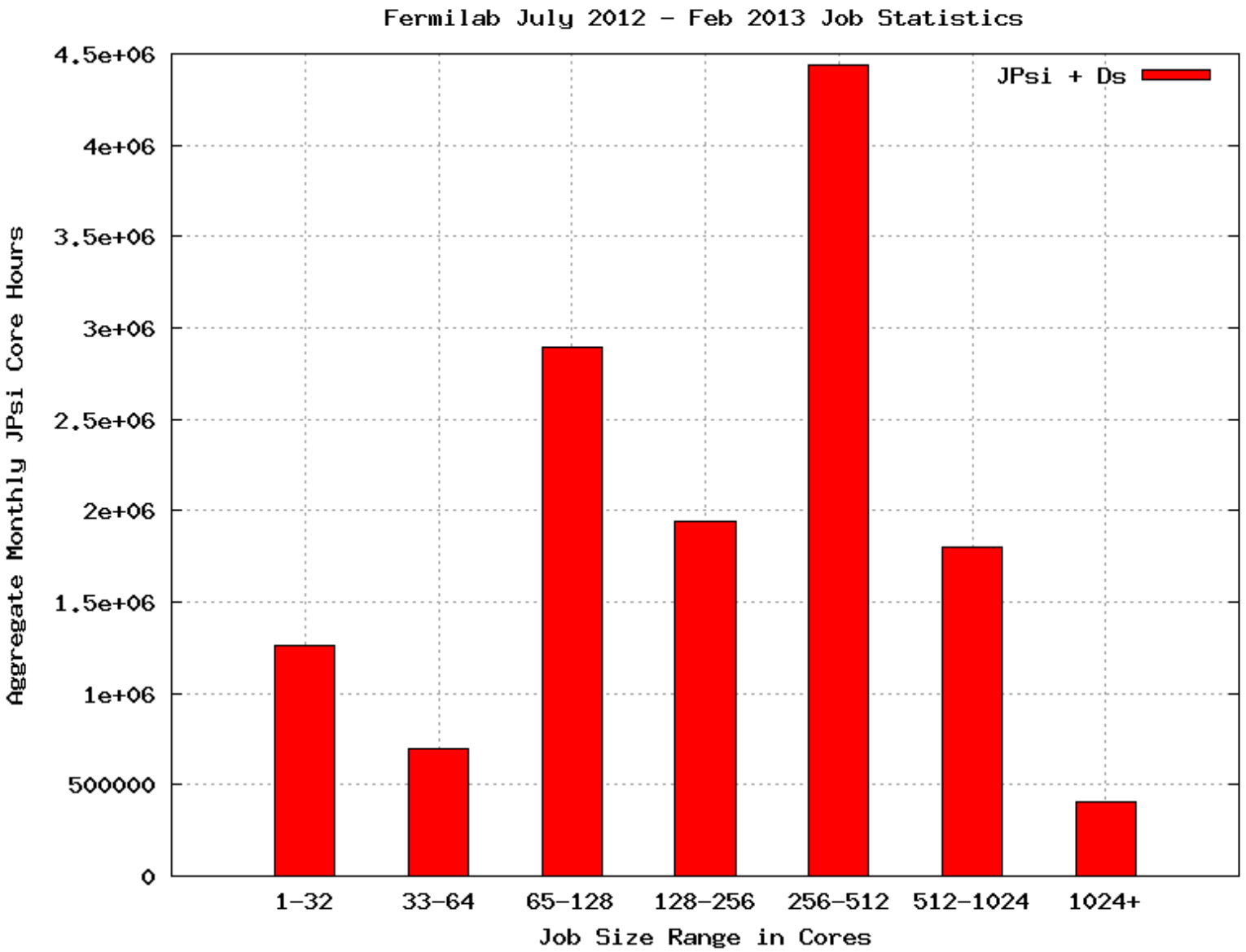


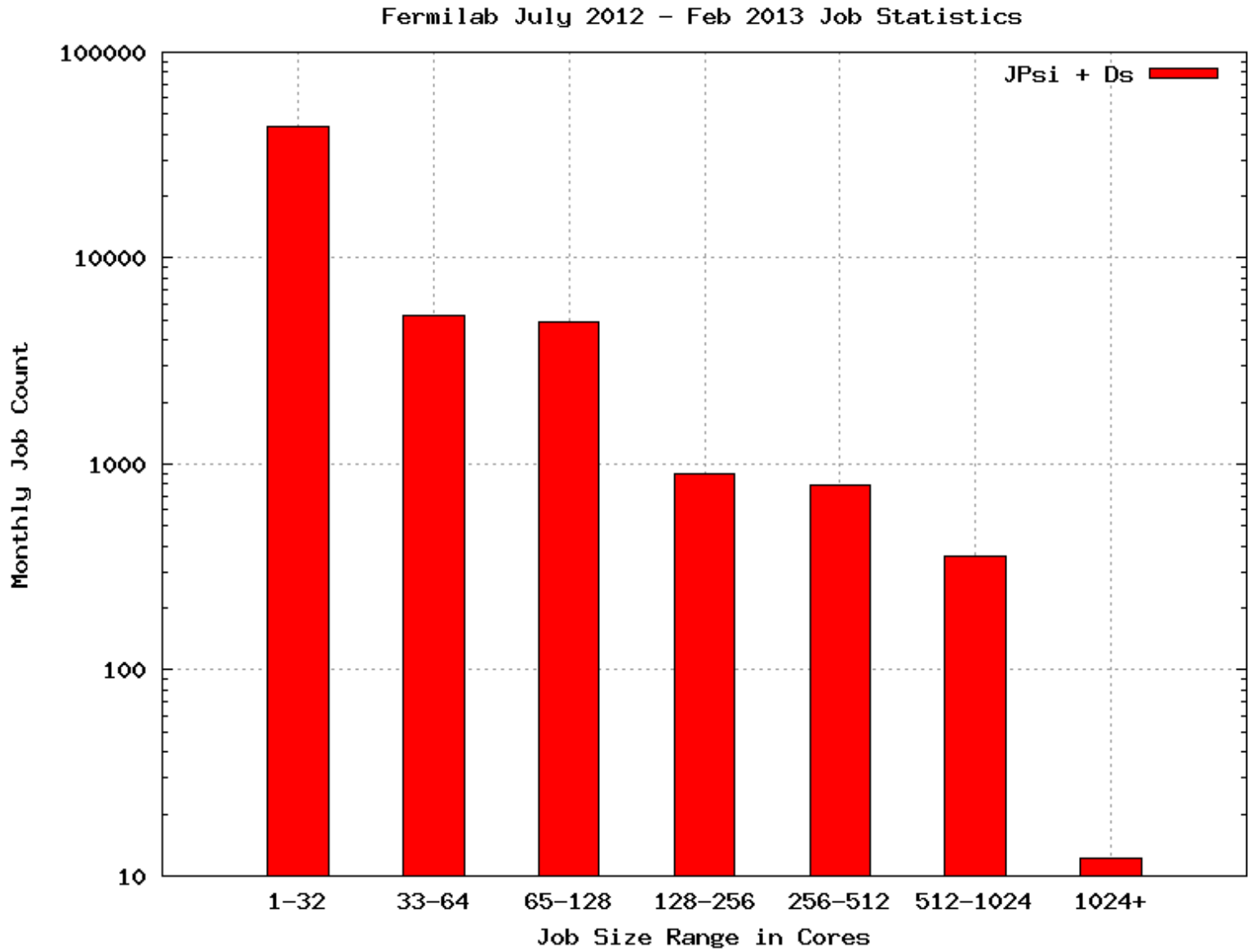
- Utilization of /lqcdproj will always increase to fill all space. This is a good thing (disk is expensive – we don't mind you using it).
- But:
  - Lustre misbehaves when the pools get above 95% fill. Please be responsive to our requests to clear space.
  - For our planning purposes, it is critical that in your proposals that storage requests are reasonably (factor of 2) accurate. We have had instances of both large overruns and under-utilization. We can adjust budgets annually, but we need reliable data.
  - We have seen again this year I/O patterns that occasionally saturate our Infiniband network or the aggregate bandwidth of the 114 pools
    - We have put new tools into place that allow us to characterize I/O bandwidth of nodes and jobs
    - It is helpful to us if your applications log I/O times and/or rates, and you let us know when your jobs are adversely affected by congestion

Statistics

- Since April 1, 2012, including JPsi, Ds, Dsg
  - 986K jobs (1.25M including Kaon)
  - 211.0M JPsi-core-hours
  - We did not charge for Kaon (an additional 4.0M JPsi-core-hours)
  - 1137 GPU-KHrs
- USQCD users submitting jobs:
  - FY10: 56
  - FY11: 64
  - FY12: 59
  - FY13 to date: 54


Fermilab LQCD Cluster FY12-FY13 Utilization
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
Percent Utilization
JPsi
Ds
Dsg
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
June
July
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Month
The dip from June through September was noted at the DOE. It is very important that USQCD have high utilization of our dedicated hardware.


Fermilab July 2012 - Feb 2013 Job Statistics
JPsi + Ds
Aggregate Monthly JPsi Core Hours
4.5e+06
4e+06
3.5e+06
3e+06
2.5e+06
2e+06
1.5e+06
1e+06
500000
0
1-32
33-64
65-128
128-256
256-512
512-1024
1024+
Job Size Range in Cores


Fermilab July 2012 - Feb 2013 Job Statistics
JPsi + Ds
Monthly Job Count
100000
10000
1000
100
10
1-32
33-64
65-128
128-256
256-512
512-1024
1024+
Job Size Range in Cores

Progress Against Allocations

- Total Fermilab allocation: 174.10M JPsi core-hrs
  1092.3 GPU-KHrs
- Delivered to date: 157.5M (90.4%, at 79% of the year)
  834.4 GPU-KHrs (76.4%, at 79%)
  - Does not include disk and tape utilization (11.25M + 0.7M)
    - 500 TiB of disk, 238 LTO-4 equivalent new tapes
  - Does not include 2.93M delivered without charge on Kaon
  - Class A (18 total): 8 finished, 3 at or above pace
  - Class B (5 total): 0 finished, 0 at or above pace
  - Class C: 2 for GPUs, 6 for conventional

Summer Loadsheds

- During very hot summer days, we must throttle back or power down some of our machines because inadequate cooling is available
- Last year, we had load sheds May 25-29, June 15, June 28-29, July 3-9, and July 17-18
- Removal of a beam-line berm near our outside condensers lowered the fraction of load shed required compared to 2011
- We lost a total of approximately 2.8% of our nominal annual computing capacity. Since we strive to achieve about 92% uptime, this loss was well within the 8% downtime allocation.
- The Bc cluster will be housed in a computer room that will not be subject to load sheds this summer
- The lab anticipates funding in FY14 to do infrastructure improvements, including remediating these cooling issues. However, we will still have to deal with load sheds this summer.

User Support

Fermilab points of contact:

- Best choice: lqcd-admin@fnal.gov
- Don Holmgren, djholm@fnal.gov
- Amitoj Singh, amitoj@fnal.gov
- Jim Simone, simone@fnal.gov
- Ken Schumacher, kschu@fnal.gov
- Rick van Conant, vanconant@fnal.gov
- Alex Kulyavtsev, aik@fnal.gov
- Yujun Wu, yujun@fnal.gov
- Paul Mackenzie, mackenzie@fnal.gov
- Please use lqcd-admin@fnal.gov for requests and problems

Future Facilities

- Questions posed by the conveners, addressing the time frame of the next facility proposal (FY15-FY19):
  1. What are FNAL's views on future LQCD computing facilities?
  2. What are the priorities for FNAL computing?
  3. How are we influenced by users' computational requirements?

FNAL's Views on Future Computing Facilities

- Caveat: these may be more Don's views than FNAL's views (since I'm writing this the night before the meeting)
- FNAL is very interested in continuing to host clusters and accelerated clusters.
  - With Run 2 over and LHC pausing, for the next several years LQCD has a very high profile at FNAL and consequently access to good computing facility spaces and support (networking, storage).
  - FNAL computing management is definitely committed to the lab excelling as a host facility.
- For at least the next several years, a mixture of conventional Infiniband cluster and accelerated (NVIDIA, ATI, Phi) clusters will be needed to match user demand.
- Storage will continue to become a more significant bottleneck (both storage volume, and I/O performance – both disk and tape). Tape is (unfortunately) not going away. We need better information from the systems we operate, and from the user community. We need to understand future costs better.

FNAL Priorities

- Top priority is meeting the computing requirements of USQCD analysis computing
- Leverage FNAL capabilities
  - Cluster expertise (reconstruction farms and real time triggers)
  - Storage (networks, file systems, data movement)
- LQCD hardware often provides design ideas and prototyping that is useful to other programs at the lab
  - e.g., for several years Ds-type machines (quad-socket Opterons) have been the standard used for Run 2, FermiGrid, CMS Tier 1
  - other programs at the lab are now becoming very interested in GPUs and Intel MIC architecture

How are we influenced by users' computational requirements?

- Project level:
  - DOE funding comes in ~ 5-year cycles. At the beginning of each cycle (e.g., now) we provide input to the Executive Committee on hardware capabilities and projected costs. Capabilities are defined in terms appropriate to meeting the computational requirements of LQCD (*e.g.* sustained TF). This information influences the scientific reach in the funding proposal. Once approved, this also sets the baseline for the budget and for how our level of success is measured.
  - Annually, the program and site managers work with the EC to determine how the budget is distributed among the three labs and what types of hardware will be purchased. This is driven by scientific needs and the result should optimize the portfolio of computing facilities operated by the hardware project. A fairly rigorous argument is made to the DOE and approved by the NP and HEP program managers. This year's Alternative Analysis was particularly complicated (mixture of BG/Q at BNL and a cluster at FNAL), and complexity of this task has increased greatly with the introduction of GPU accelerators.

How are we influenced by users' computational requirements?

- Site (FNAL) level:
  - When appropriate, we interact with the principal users of the major allocations at FNAL to understand requirements in depth. Prior to GPUs, design of clusters was straightforward – maximize memory bandwidth per dollar – and other than input on memory per processor or storage space we did not seek much guidance. For the Dsg cluster, we sought a lot of guidance from the EC, PI's, and experts, and successfully (I believe) delivered a GPU cluster that runs large parallel jobs (32+ GPUs) well. We will need considerable input for the FY14 machines.
  - We perhaps are more influenced by requirements after a procurement, so the next machine sometimes seems to fix those requirements unfulfilled by the last cluster
    - E.g., we've been playing catch-up on storage requirements for some time
    - Our communications are with USQCD users of the FNAL machines. Perhaps we need to find a way to have broader input from USQCD (since the same groups of users have tended to have allocations at the same sites from year to year). Let's give this question to the full USQCD community.
  - I've been concerned for some time that we (project-wide) are too influenced by the availability of new hardware/software and not enough by the requirements of the running projects. Since we buy new hardware every year, and since USQCD has access to many other facilities, delaying the purchase of the latest and greatest hardware does not carry that significant a penalty.

Questions?