Jefferson Lab LQCD Computing Status & Plans

Chip Watson

Scientific Computing

Sandy Philpott, SciComp Operations Manager

Outline

- Resources at JLab
  - compute: CPU & GPU
  - disk, tape, and networking
- Banking Core Hours / Load Balancing
  - now able to share resources with experimental physics!
- Operations Summary
- Lustre Upgrade & Partitioning
- Plans for next 2+ years
  - decommission ARRA clusters over next 2 years
  - install a new machine Spring/Summer 2016…

Jefferson Lab Resources

✧ Conventional Clusters

2012 **12s** 276 nodes 4416 cores

16 cores, 32 GB per node, hyperthreading on (Sandy bridge CPU)

full fat tree QDR fabric with no overscription, supporting jobs of up to 4K cores, although in practice no one is running above 1024

2009-2010 **9q, 10q** 544 nodes 4352 cores

8 cores, 24 GB per node, hyperthreading on

partitioned into 13 * 32 node + 1 * 128 node partitions, no oversubscription

Jefferson Lab Resources

✧ GPU Clusters

2012 **12k** 42 nodes 168 GPUs 4.7 GB / gpu

host: 16 cores, 128 GB per node, hyperthreading on, full FDR IB

2009-2010 **9g, 10g** 114 nodes 456 GPUs 1.5-2.7 GB / gpu

host: 8 cores, 48 GB per node, hyperthreading on

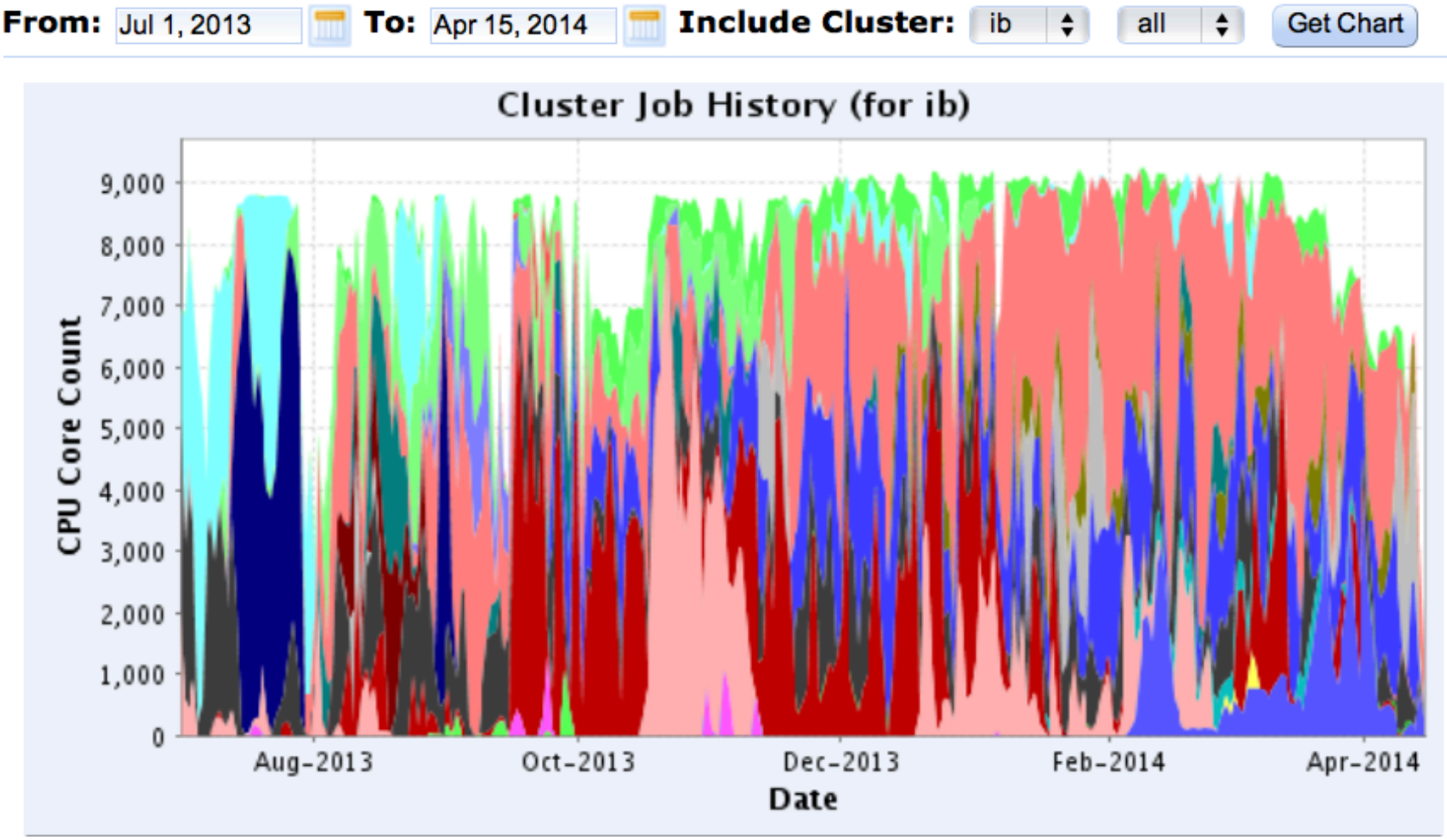
70% in 3 generations of gaming cards:
GTX-285, GTX-480 / GTX-580, GTX-690

30% Tesla 2050, “Fermi” GPU, 2.7 GB / gpu

Infiniband: mix of ½ SDR (i.e. in half bandwidth slot), ½ DDR, ½ QDR, full QDR for GTX-690 (dual GPU cards, leaving full bandwidth slots open)

GTX-285’s will be retired June 30, GTX-480’s at most a year later

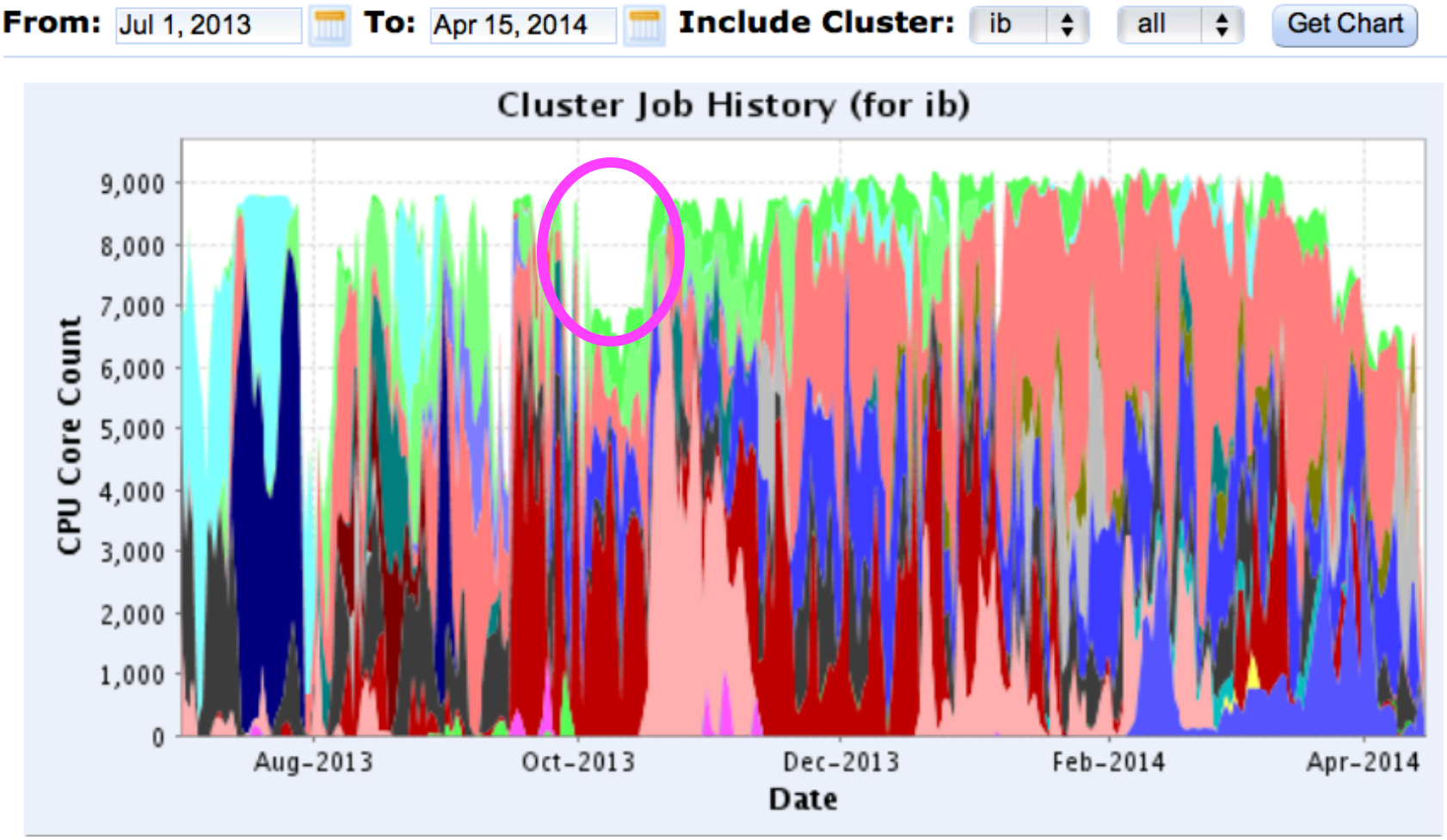
Conventional Resources



Job utilization since July 1, 2014

(all graphs available online at lqcd.jlab.org)

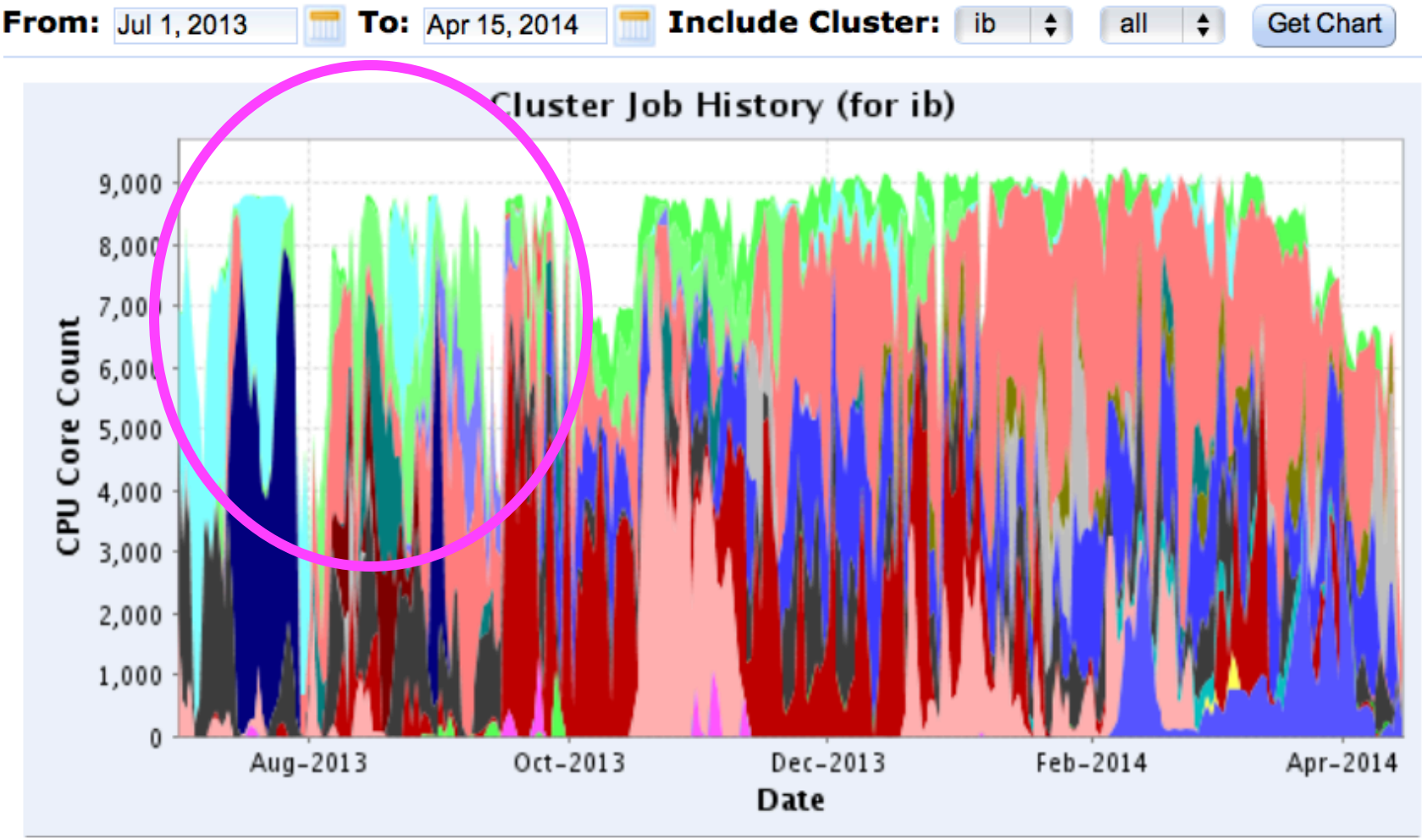
Conventional Resources



Reduced operation during pending gov't shutdown, ½ M lost

(budget problems also caused us to reduce staff)

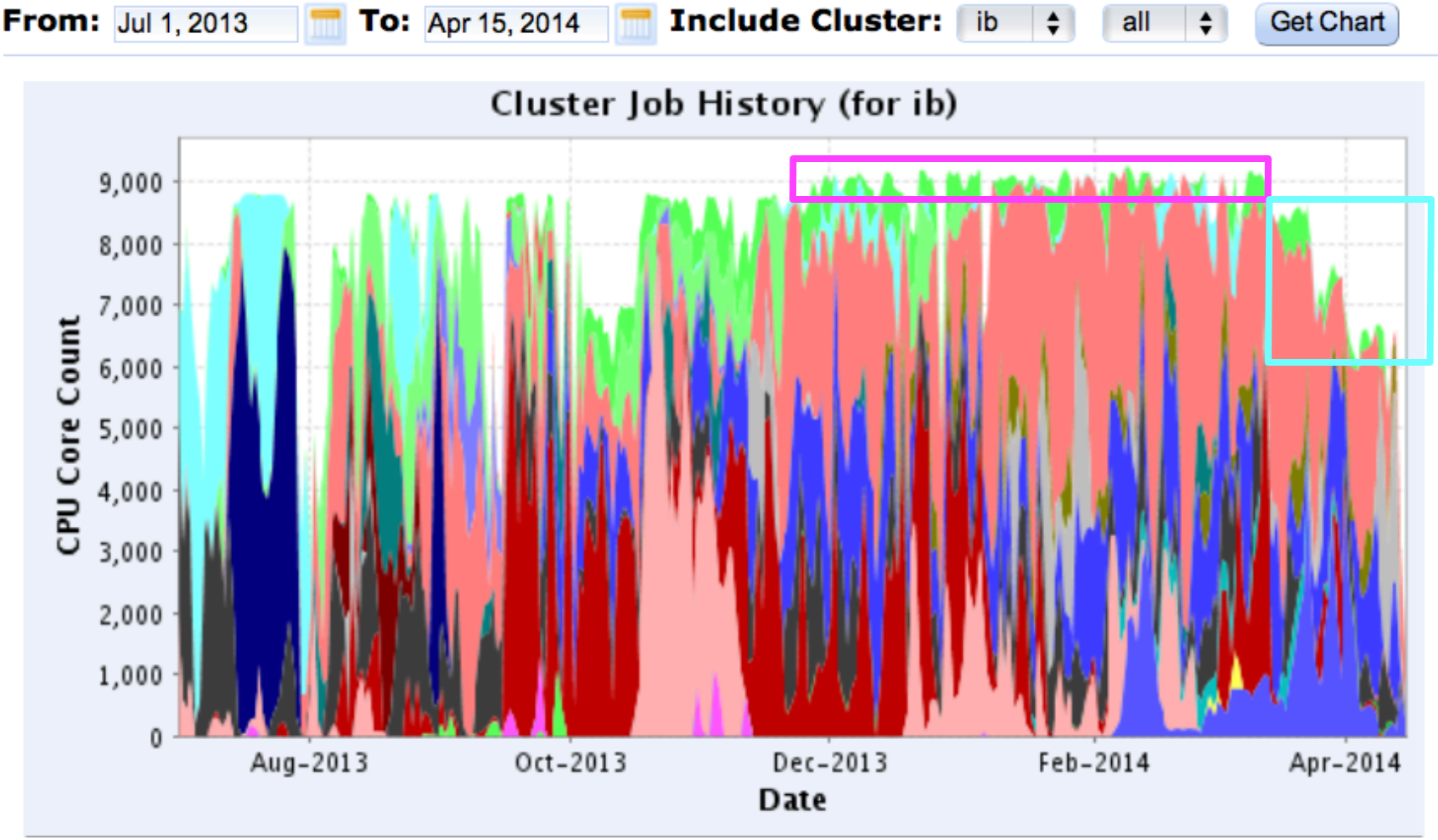
Conventional Resources



Underutilization in July-September, ~ 2M core hours lost

(yes, this is only a 1.5% effect, but larger than some entire projects)

Conventional Resources



two other interesting features this year…

Resource Sharing (banking hours)

We decided to deal with the annual under-utilization problem via a simple software fix…

Goals:

- never waste core hours, even when users are otherwise busy
- support GlueX & CLAS-12 data challenges (using LQCD cores)

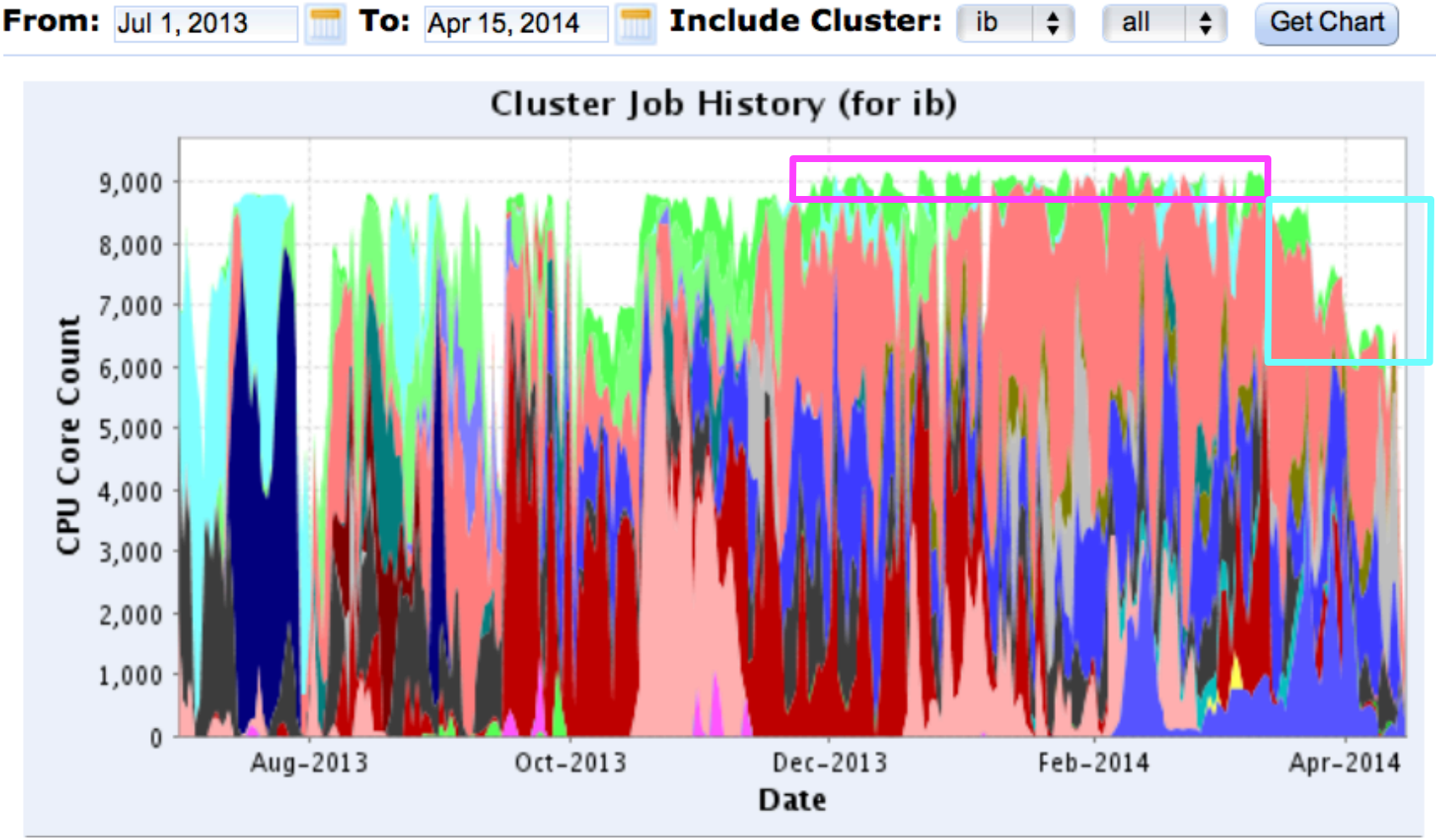
⇒ Move capacity between LQCD and Experimental Physics

Experimental Nuclear Physics bought nodes similar to our 12s purchas (identical nodes except no QDR) plus purchased additional Ivy Bridge nodes in the fall of 2013. They loaned LQCD 512 cores for 3½ months to build up a credit.

Total banked was almost 1 million core hours. (Thanks to all who suffered a little pain getting the physics nodes running in our environment.)

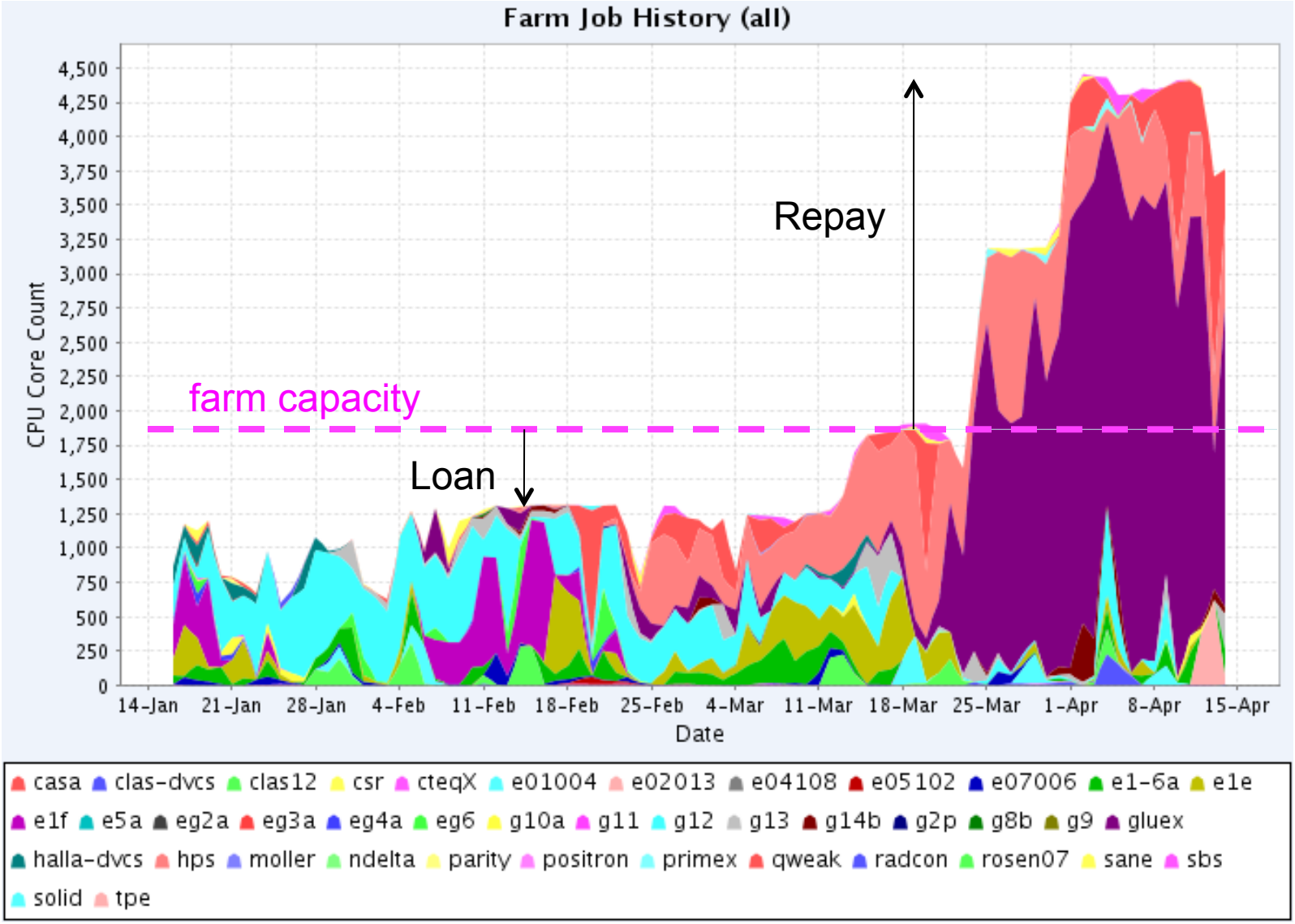
The return flow was 4x larger, 2048 cores of 12s, thus about 22% of our capacity, but sufficient to increase the Physics "farm" 2.5x …

Conventional Resources



1M core hour loan from physics……returned for data challenge

Viewed from the farm (20 per 16 cores)



Win-win

Future Impacts

- As long as the farm is busy, LQCD users don't have to worry as much about losing any of their allocations due to wasted core hours
- Physics can do data challenges at much larger scale than their resources for this year and next (delaying purchases until needed for production running)

We anticipate similar scale data challenges late summer (when LQCD typically lags) and again next winter (just after a farm upgrade).

The swings might grow larger as we try to use LQCD as a flywheel for the peaks and valleys of commissioning and early running of the GlueX detector. (We will work with users to ensure this won't harm LQCD productivity.)

Operations: Conventional

USQCD Project Allocation Usage (13-14)

(Core hours for each cluster are converted to 13-14 hours based upon measured relative performance)

Project Name	Proj. Allocation	Proj. Used Hours	Annual Pace	Monthly Pace	Hour Remaining	Overused
NPLQCD	41,060,000	30,492,247	94%	105%	10,567,753	0
Spectrum	36,620,000	33,315,924	115%	40%	3,304,076	0
thermo	21,610,000	16,322,479	95%	88%	5,287,521	0
Higgs	18,320,000	12,932,140	89%	126%	5,387,860	0
Structure	5,400,000	5,826,795	136%	0%	0	426,795
TMD	3,830,000	3,459,111	114%	639%	370,889	0
isoClover	150,001	152,080	128%	0%	0	2,079
EigenSpect	20,001	13,759	87%	0%	6,242	0
SMhisq	20,001	10,117	64%	0%	9,884	0
emc	10,001	6,117	77%	3%	3,884	0
Total	127,075,011	102,530,768	102%	98%	24,973,117	428,874

- Systems are running extremely well
- Utilization high other than summer slump
- On track to deliver all allocations by June
- Monthly pace is low because it includes the 2 week data challenge

Operations: GPU

(GPU Hours are converted to allocation hours based upon measured relative performance)

Project Name	Proj. Allocation	Proj. GPU Hours	Annual Pace	Monthly Pace	Hour Remaining	Overused
Spectrumg	3,028,000	2,295,773	96%	43%	732,227	0
thermog	2,331,000	1,832,599	99%	94%	498,401	0
Higgsg	742,000	569,689	97%	225%	172,311	0
Structureg	410,000	458,376	141%	340%	0	48,376
NPLQCDg	187,000	157,168	106%	149%	29,832	0
chiQCDg	90,001	75,827	106%	0%	14,174	0
emcg	5,001	0	0%	0%	5,001	0
Total	6,793,002	5,389,431	100%	101%	1,451,947	48,376

- All allocations are on pace to finish in June
- Despite the complexity of the system (many GPU flavors), the utilization has been very high for all GPU types
- K20s are quite popular (over subscribed)
  - good performance, especially in double precision
  - large host memory (128 GB) allows most jobs to fit into one node
  - cost per node is 4.4x conventional node, more than worthwhile for certain applications

Jefferson Lab Storage & Networking

✧ Disk Systems

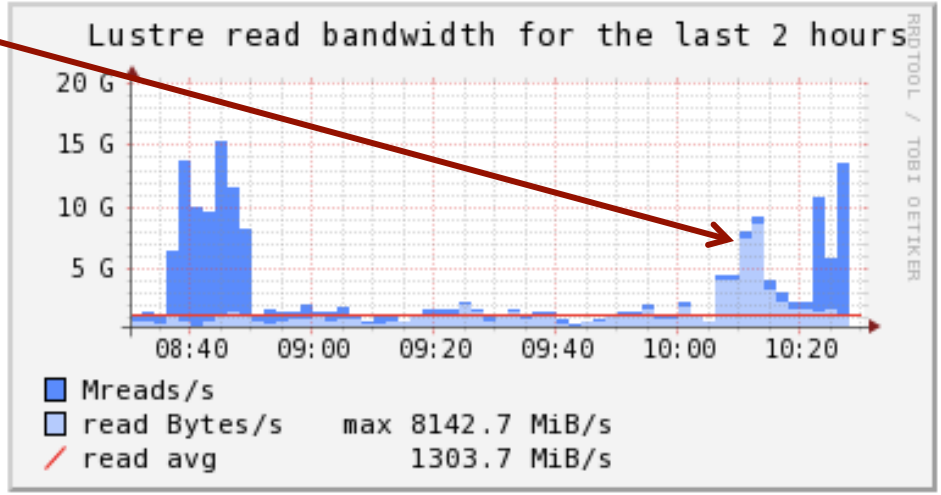
Lustre 896 TB, 640 TB allocatable (target: 80% full, peaks up to 90%, auto managed)

/cache/project…	420 TB	write-through cache to tape & delete to stay under quota
/volatile/project…	220 TB	delete to stay under quota

8.1 GB / sec aggregate bandwidth

this will soon be increased ~20%

100 MB/s – 1 GB/s single stream



ZFS servers 62 TB

/work/project	62 TB not backed up
/home/…	1 TB nightly backup

✧ Tape Library

14 frames; can add 1 or 2 more when needed (1300 slots each, 2.5 TB / slot) and can upgrade low density frames to high density, or buy 2nd library

2 LTO-6 drives, 10 LTO-5 drives, 2 LTO-4 drives => 1.6 GB/sec to / from tape

✧ Offsite Gateway

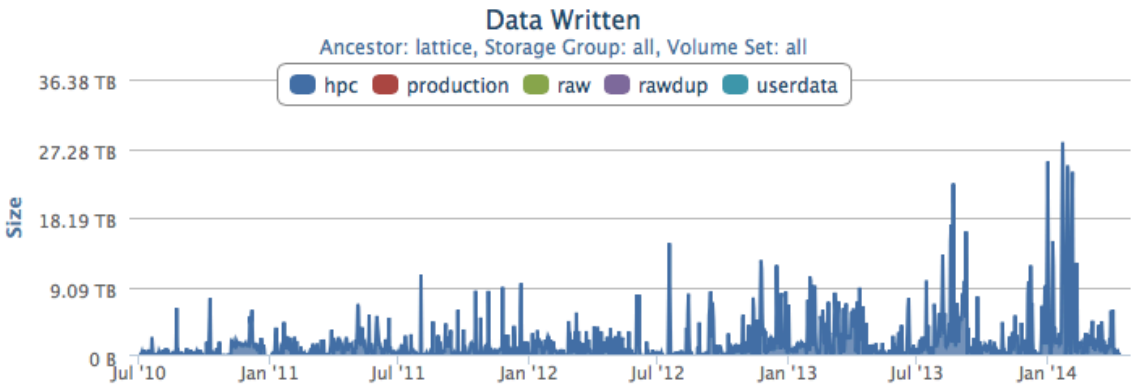
Globus Online working well for large bulk data transfers, up to around 5 Gbps

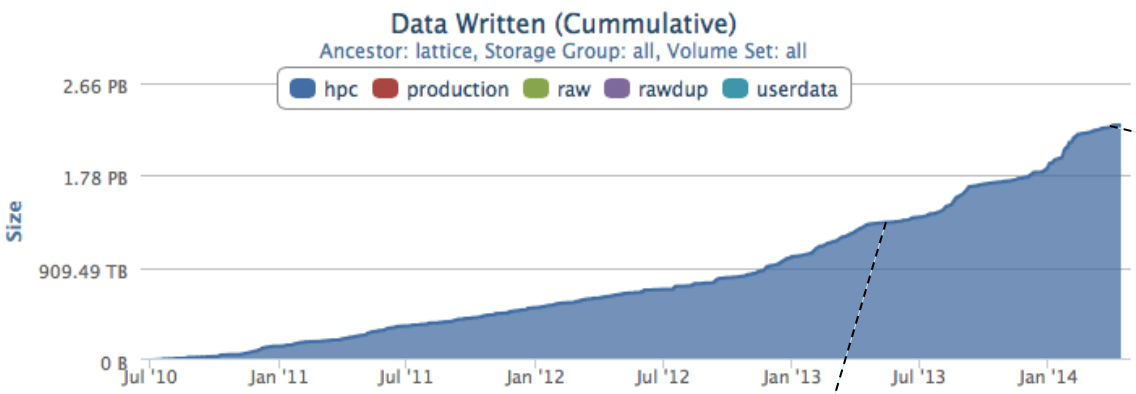
Restricted shell implemented on the offsite gateway host (same as JLab login hosts)

Storage Evolution

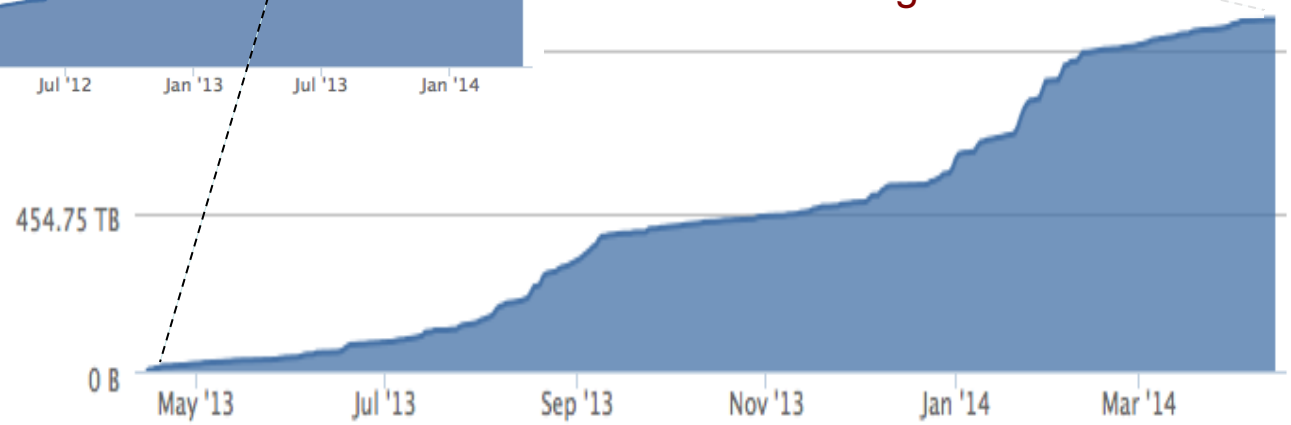
- **Lustre Upgrade and Partitioning**
  - Each generation of server is higher capacity, higher performance (similar in units of MB/s per TB)
  - Rolling upgrades: retire 5 year old systems, maintain capacity
  - New version of Lustre, and partitioning by performance
    - Prototype running Lustre with 2 partitions: fastest / newest, and older / slower
    - This will ensure that if you strip a file, all strips will be fast (or all slow)
    - Initially will use the retired 2009 servers as a testbed, along with a new higher performance MetaData Server (old one, now 5 years old, is at end of life)
    - By the end of 2014, this will be in production, with “inactive” projects moved from the main partition into the older, slower partition, freeing up highest performance disk space for active projects
  - Question: how to name this new slow partition?
    - Option 1: no new name, inactive projects on slow, active on fast, and slow ones can be accessed via links: /cache/inactiveA -> /slowcache/inactiveA
    - Option 2: give users access to both fast and slow, and let them optimize, and “charge” half as much for slower servers?

Tape Usage





- Usage increases with capacity: notable uptick in late 2012
- Varies depending upon job mix, phase of projects (below)
- Currently running at over 1PB / year, at rates up to 3x higher
- Cost per year: media $22K, slots (JLab cost) $12K
- Future cost to project will climb in 2015, as we go to higher density tapes to save slots, then shrink again in 2016+.



Scientific Computing Home

SciComp

Home

Search this site:

Search

watson

- Books
- My account
- weblinks
- Create content
- Frequently Asked Questions
- Content summary
- Feed aggregator
- Administer
- Log out

Go

Enter the username to masquerade as.
Quick switches:

- drupaladmin
- philpott

LQCD User's Guide

View | Edit | Outline | Revisions | Track | Devel

This guide is meant to cover everything related to using the LQCD Computing Facilities at Jefferson Lab, from batch systems to file storage and management. More detailed information on the hardware available at Jefferson Lab is contained in the companion book, Scientific Computing Resources.

New users should start with the first chapter, which gives an overview of what you will need to know. Then you can skip to whichever chapter you need. The menu on the right will make it easy to skip around. Examples of how to use the system are contained in several of the chapters, and additional useful tips are contained in the Frequently Asked Questions chapter at the end.

- Getting Started
- Batch System
- MPI
- Using GPUs
- Using MIC cards
- File Systems
- FAQ

Getting Started ›

LQCD User's Guide

- Getting Started
- Batch System
- MPI
- Using GPUs
- Using MIC cards
- File Systems
- FAQ

New Drupal web pages organized as “books”. Suggestions for additional material welcome! Click on “New users start here” from lqcd.jlab.org.

Future Plans

2015-2016

Operate current resources (minus oldest gaming cards): run the late Fall 2009 clusters through June 2015, and mid 2010 clusters through June 2016. Longer than usual due to absence of hardware for 2015. Experimental Physics grows to match the size of LQCD, enabling efficient load balancing (with careful balance tracking).

2016

Install new resource of scale ~$1.8M using funds for 2 fiscal years (16 & 17) to achieve same scale as recent procurements.

FY2016 – FY2017

JLab will be the deployment site for the first cluster of LQCD-ext II. This resource will be installed in the current location of the 9q / 10q clusters (same power and cooling, thus lower installation costs).

Architectures To Watch

- Intel Xeon "Skylake" 14nm
  - DDR4, AVX3.x, PCIe4, 18+ cores, … (i.e. better, faster x86-64)
- Nvidia Tesla "Pascal"
  - stacked memory for higher local bandwidth
  - custom network connection to address PCI bottleneck
- Intel Xeon Phi "Knights Landing"
  - 72 cores
  - self hosted (not PCI card, minimal Amdahl's Law effect)
  - 8-16 GB stacked memory plus 6 channels of DDR4 (up to 384 GB)
  - Infiniband or Cray HPC network connection (version –F)

Both advanced architectures will have higher memory bandwidth stacked memories, and will depend upon optimized data structures.

FYI, for Xeon Phi you can start now on the 16 node quad Xeon Phi (Knights Corner) **12m** cluster. Just request a class C allocation.