JLab Status

April 2017 All Hands Meeting

Chip Watson, Sandy Philpott
Jefferson Lab

Outline

SciPhi XVI cluster (16p, knights landing)

File System Upgrades

Operations

Jefferson Lab
Thomas Jefferson National Accelerator Facility

16p   Top500 #397   Green500 #10

Xeon Phi 7230, 64 cores/node

192 GB memory +

16 GB high b/w memory

100 Gbps OmniPath

1 TB disk

380 M Jpsi core hours/year

50 TB memory in 256 nodes

4 racks of 64 nodes, 264 total

Integration (and 5th rack)

− Extra 8 nodes (4 IT div),

− 4 LNET routers
  OP to QDR IB

− 2 IP routers
  OP to QDR IB

OP fabric: (48 port switches)

- 8 leaf switches: 32 nodes, 16 uplinks

- 3 core switches

    (w/ extra nodes, routers)

- mix of copper & fiber cabling

Jefferson Lab

Newest Machine: " **SciPhi XVI** *"*

Scientific Computing, Xeon **Phi** **(16p)**

Why KNL?

Low cost, high performance node (cost less than dual socket)

- AVX512 double width vector units

- On package, high bandwidth memory

➢ 25% **higher bandwidth/$** than pure P100 (not considered) and 90% higher than 50:50 P100:x86

➢ 18% **higher dslash/$** (32^4) than 50:50
  > 450 GFlops dslash on a single node

➢ 80% **higher zgemm/$** than 50:50

➢ 2x - 5x **better USQCD code coverage**

➢ Best able to accommodate large memory footprint jobs (low cost, moderately high memory nodes)

JSA

# Scientific Computing Network Fabrics
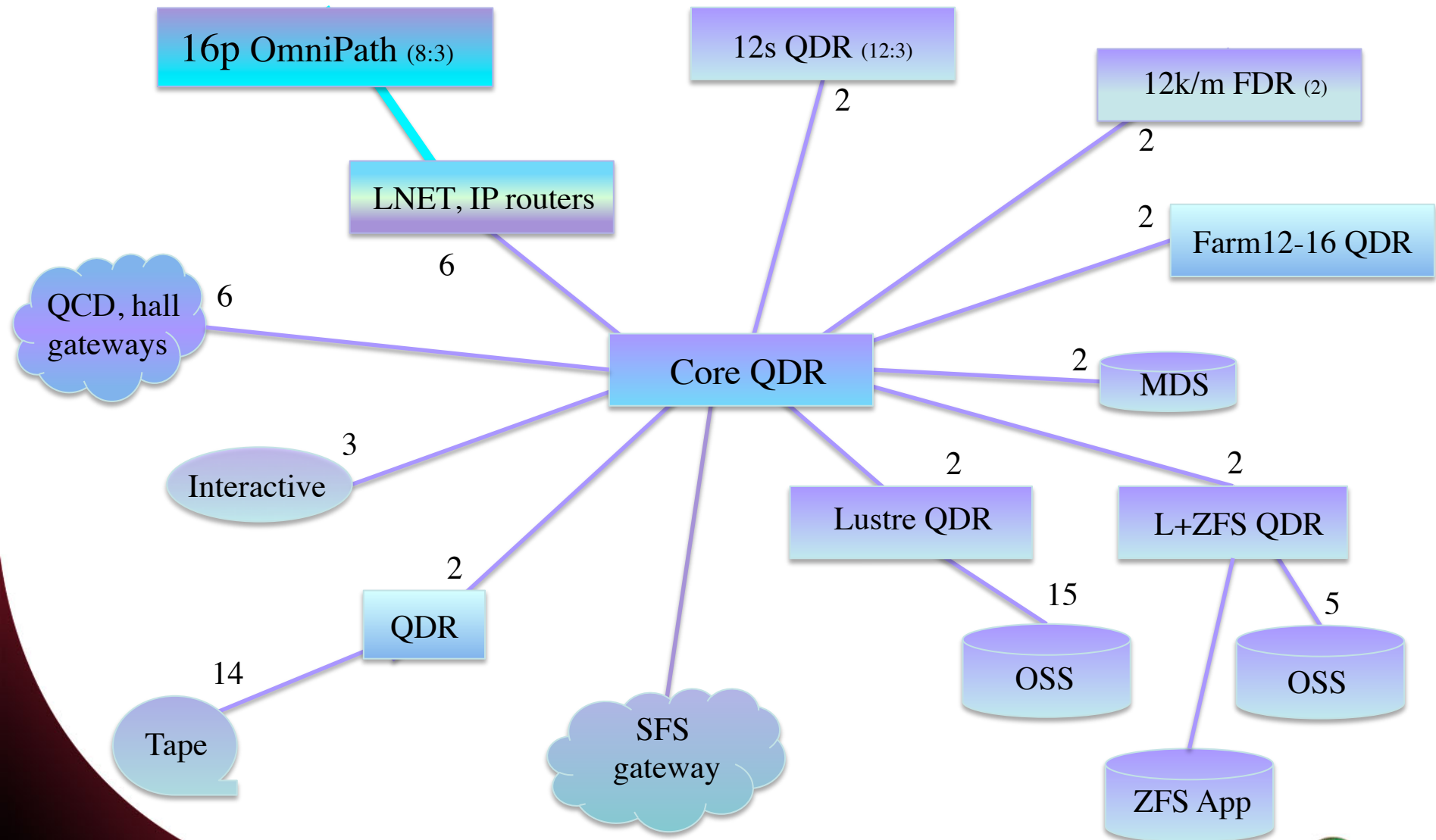
Update: KNL weaknesses

- Can't get full network bandwidth with 1 thread or core (we knew this at award); 4 cores is good, 8 is very good (see Bob's talk tomorrow for recent "grid" benchmarks)

- Codes that fit in the 16 GB of high bandwidth Multi-Channel DRAM on-package memory run very fast, but if not…

- Code not written to manage memory in a sophisticated way can use MCDRAM as an L3 cache… this works, but requires a CPU reset to change modes

  - Flipping CPU modes in not entirely robust (yet), and so is now done manually by system administrators

- Performance in cache mode degrades over time! CPU reset restores performance after a short while

- Failure rate on our early release chips has been higher than expected, about 1 / week, almost all due to the MCDRAM

16p Strengths

- Cache mode allows running some standard codes w/ no code modifications
- Larger memory per node -- 192 GB; this allows many jobs to run single node (more than 90% at JLab has been single node, see graph)

Core Hours vs. Job Size since Oct 1
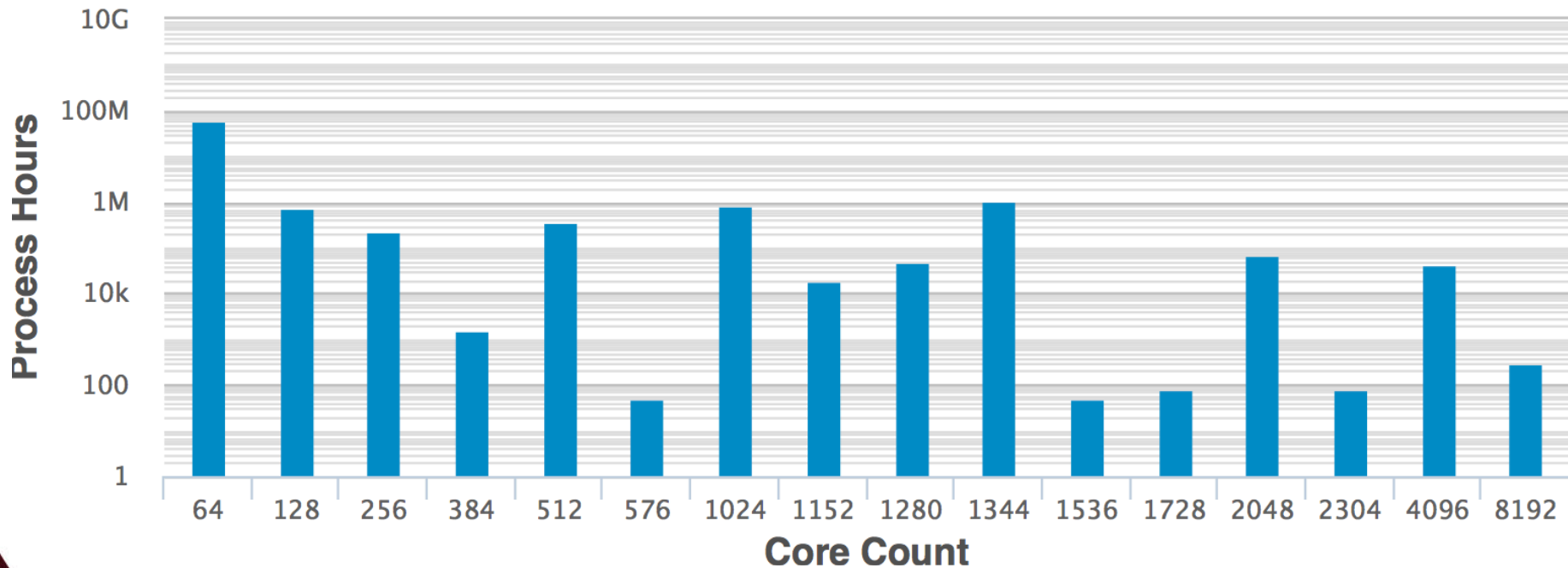
Update: knl evolution

- There are already newer bios and OmniPath software versions, and when convenient, we will upgrade…

- 4 test nodes are available via the testphi queue with

    CentOS 7.3 (we are at 7.2 since October)

    xppsl 1.5.1 (current: 1.4.1)

    OPA 10.3 (current: 10.1)

- Intel is working on even newer bios that might solve both the reboot / CPU mode change unreliability, and the MCDRAM performance degradation (fragmentation?)

Jefferson Lab

JSA

Operations

JLab Resources: 3 Architectures

1 Xeon Phi 7230 (KNL) cluster, 264 nodes, 16.9k cores

1 GPU cluster (k20), 45 nodes, 180 GPUs

1 Infiniband cluster, 4.3k cores

2.4 PB Lustre file system
- shared with Experimental Physics, 55% LQCD
- 20 servers
- >10 GB/s aggregate bandwidth (have yet to see a saturation)
- Auto managed to < 80% full to avoid fragmentation

12 PB tape library, shared, ~ 3 PB LQCD
- LQCD growing at around 40 TB / month
 much lower than requests would have implied

New Web App

As announced last year as "in progress", all of our web pages have now been updated to become a new web app

- More responsive, changes in table sorting persists as you jump from one page to another and back inside the web app

- Fluid layout, good for cell phone, tablet or large monitor

- Better auto-scaling of graphs, tables (still being tweaked)

- More operations information available to both LQCD and ENP

- All graphs/tables in this presentation come from the new app. If you have any great ideas for new views of the data, email me your suggestions!

> » Jie Chen, Ying Chen, Chris Larrieu developers
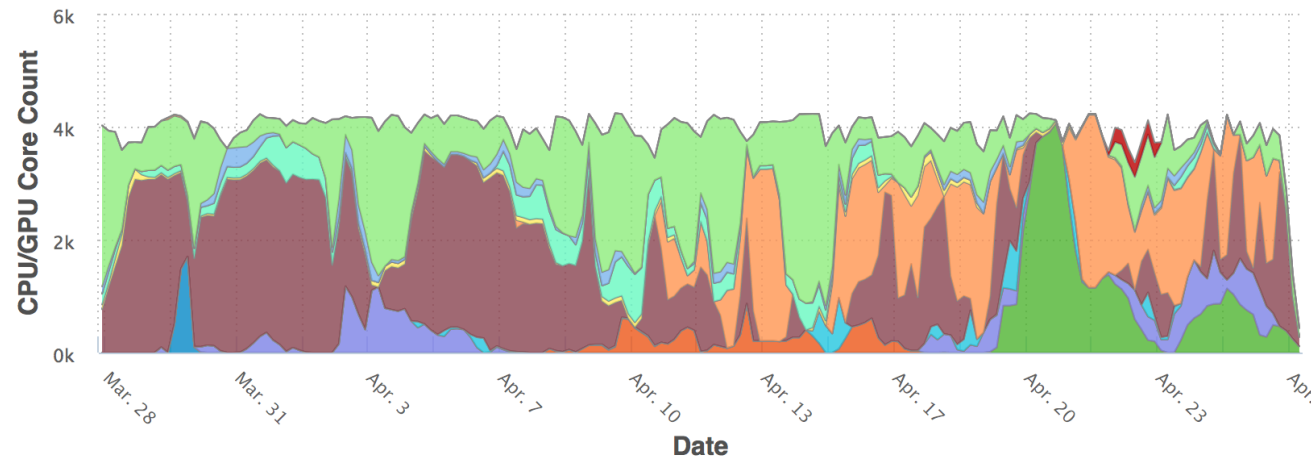
Jefferson Lab

Operations & Utilization

Completed Job History (16p nodes)
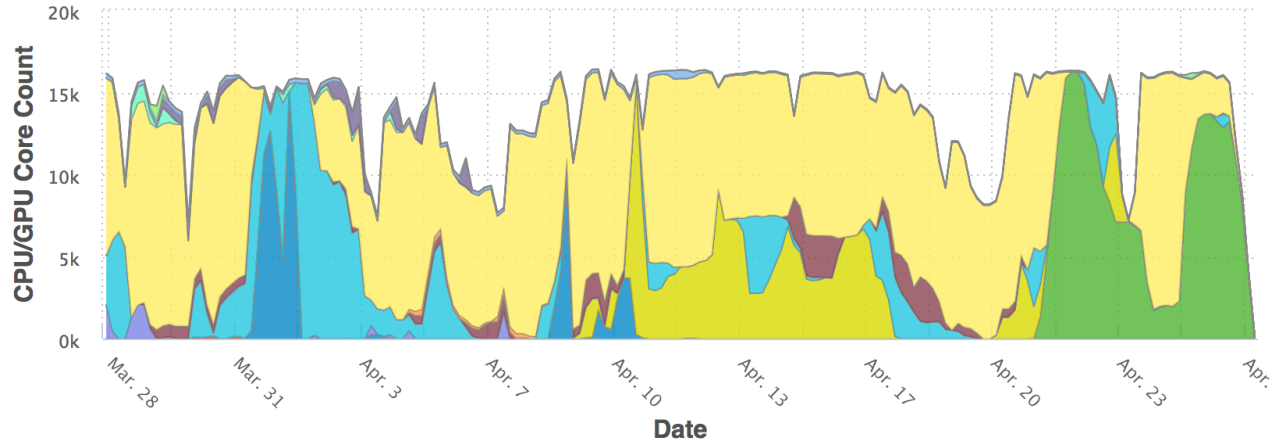
Overall, LQCD running well, with occasional dips in utilization on 16p (when queues empty)

GPU utilization (not shown) is very low for this month; allocations are mostly done

Colors are different USQCD projects/users

Allocations Progress (mixed)

Project Name	Allocation* ▾	Project Used Hours**	Annual Pace	Monthly Pace	Remaining Hours	Overused
Spectrump	66,000,000	99,664,198	198%	118%	0	33,664,198
thermop	31,000,000	92,536,070	391%	528%	0	61,536,070
Delta	26,300,000	3,901,789	19%	117%	22,398,210	0
muon	22,500,000	1,935,649	11%	11%	20,564,350	0
NPLQCDp	15,600,000	6,722,652	56%	0%	8,877,347	0
gMinus2	15,000,000	377,917	3%	0%	14,622,082	0
qcdqedta	11,300,000	3,046	0%	0%	11,296,953	0
cedmpffp	10,500,000	611,558	7%	50%	9,888,441	0
fermimilcheavylight	7,500,000	3	0%	0%	7,499,996	0
rhqbbarp	6,600,000	4,682,185	93%	42%	1,917,814	0
	212,300,000	**210,435,067**	**130%**	**133%**	**97,065,193**	**95,200,268**

(above) 2 largest knl allocations have well overrun.   3rd is ahead on monthly pace

Many have not yet reached full running pace.

Project Name	Allocation*	Proj. GPU Hours**	Annual Pace	Monthly Pace	Remaining Hours	Overused
Spectrumg	1,425,000	1,944,986	166%	44%	0	519,986
NPLQCDg	1,018,000	430,144	51%	0%	587,855	0
	2,443,000	**2,375,130**	**118%**	**26%**	**587,855**	**519,986**

GPU hours already exceeds plan for the year due to late turn-off of gaming cards (again).

Jefferson Lab

IB Allocations Progress (mostly good)

Project Name	Allocation*	Project Used Hours	Annual Pace	Monthly Pace	Remaining Hours	Overused
Spectrum	16,470,000	17,157,556	126%	36%	0	687,556
btokpi	16,310,000	14,751,106	110%	172%	1,558,893	0
thermo	8,200,000	12,440,870	184%	60%	0	4,240,870
NPLQCD	6,660,000	761,802	13%	0%	5,898,197	0
cedmpff	6,320,000	5,473,739	105%	493%	846,260	0
StochLapH	5,800,000	3,971,011	83%	130%	1,828,988	0
chiQCD	5,350,000	72,333	1%	0%	5,277,666	0
rhqbbar	4,070,000	3,642,483	108%	5%	427,516	0
mstherm	2,900,000	1,091,286	45%	457%	1,808,713	0
fourfermi	2,460,000	0	0%	0%	2,460,000	0
PESRadius	2,000,000	3,112,365	189%	360%	0	1,112,365
su2	1,000,000	1,125,031	136%	154%	0	125,031
t-charge	36,000	619,539	2,093%	3,253%	0	583,539
JAM	10,000	0	0%	0%	10,000	0
	77,586,000	**64,219,121**	**101%**	**131%**	**20,116,233**	**6,749,361**

Conventional allocations are mostly doing well, with some starting late and other taking advantage of that to get more science done.

File Systems

Lustre /volatile and /cache

- – Upgraded to Intel Lustre last June; Intel has been very responsive to and helpful with trouble reports

- – LQCD added 512 TB new and retired 280 TB of 2010-2011 servers

- – /cache plus /volatile now managed to 1 PB

- – If Lustre is less than 80% full, pools and quotas can be overrun

- – Coming: policy changes to roll to tape old, idle projects

Next upgrade: replace ancient /home and /work

- – Dual head design, active-active, shared with ENP (they pay ½)

- – Optimized for small files, parallel make, etc. (Lustre best for large files)

- – Hybrid storage
  - ✓ SSD for ZFS-intent-log (ZIL); speeds up write acknowledge
  - ✓ SSD for read cache; augments memory cache significantly

Computer Room Upgrades (update)

To meet DOE goal of PUE of 1.4, power and cooling are being refurbished (announced 2 years ago, slow implementation)

- New 800 KW UPS (done)
- 2 new 200 KW air handlers done; refurbished 180 to be added soon

- All compute nodes have now been moved into Hot Aisle Containment (photo at left)
- All file servers, interactives, etc. will move in 11 DAYS* to dual fed power, one side of which will be generator backed (99.99% uptime)

* One day full outage planned for May 9

Documentation
(again: slowly improving; never perfect)

Drupal "book" based

– Chapter, section, subsection; next/up/back navigation

– Searchable

Three Books:

1. Scientific Computing Resources

   – Covers both shared and mission specific

2. LQCD User Guide

3. Experimental Physics User Guide

Data Management FAQ

What is the best/fastest way to move files from /volatile to /cache?

/volatile and /cache are actually remounts of /lustre/volatile and /lustre/cache. If you use the `mv` command to move files between the two and if you fully qualify the files, i.e., use `mv /lustre/volatile/PROJECT/src/files /luste/cache/PROJECT/dest/files`, then the move will be instantaneous. Doing `mv /volatile/PROJECT/src/files /cache/PROJECT/dest/files` will require a copy of all of the source files vs just a rename.

When will a file under /cache be backed up?

By default, files in /cache and 12 days old will be automatically backed up. Files with size larger than 1MB will be migrated to the tape library. A user can use srmPut to flush files to tape at anytime.

Why is it a good practice to tar small files?

Although /cache is designed to store large data files, a user can create a file with any size. However, it is recommend to pack small files into a large file if there are more than 100 files in a group. To put/get many many small files to/from the tape system might take many days, whereas a single tar file could be written or retrieved in a small fraction of that time.

Can a file/directory be renamed or moved after it is backed up?

No, a user shouldn't rename or move a file after it is put onto tape. There is a one to one mapping between the file path in cache and that in the tape library. Once a file is put into the tape library, the content, location and name should not be altered. If a file is moved/renamed, it will be treated as a new file and will be backed up again. This will not only waste tape, it will also create confusion for future users who may want to use the data. Please send a CCPR if you want to rename a cache directory after the files is backed up to tape system.

LQCD User's Guide

▷ Getting Started

▷ Batch System

▷ MPI

▷ Using GPUs

▷ Using MIC cards

▽ File Systems

  ○ /work from within a batch job

  ○ Cache manager policy

  ▷ Cache manager utilities

  ○ Data Management FAQ

  ○ Volatile management

  ○ Typical usage pattern

○ FAQ

Jefferson Lab

Questions?