Computing overview

Graham Heyes - ENP

Sandy Philpott - IT

Introduction

- Computing infrastructure for ENP is funded by the ENP division as part of the OPS budget but installed and managed by SCI group in IT. Networking and other support is provided by the CNI group in IT.
- Sandy Philpott manages the operation of the computing infrastructure (computing nodes, disk storage, tape library, etc) and will be talking about that shortly.
- The use of the ENP computing infrastructure is coordinated on a per hall basis by, Ole Hannsen (A), Dennis Waygand (B), Brad Sawatzky (C) and Mark Ito (D).
- For historical reasons the head of the data acquisition support group (me) I has several roles in the offline computing world:
  - An intermediary between ENP and IT to arbitrate, advocate and coordinate on behalf of the computing infrastructure users.
  - Gathering computing requirements and working with IT to fulfill them.
    - Technical review of plans from the SCI group - annual work plan process.
    - Managing the ENP computing budget.
  - Point of contact for computing related topics:
    - Data management plans, cyber security assessments, software quality assurance, IT steering committee, ENP rep on internal reviews.

Notable changes since the last review

- The physics computing users rely on a number of software packages, ROOT, GEANT, cernLib etc that needed better support than we were providing.
  - This time last year we formed a Physics Software Support committee.
    - Establish the roles of IT and ENP.
    - Identify which packages require support.
    - Identify who will provide the support.
    - Provide a mechanism for managing software package support.
  - Added a new IT problem reporting (CCPR) category for physics software and routed day-to-day issues to the correct maintainer - working well.
  - Started to work on centralized documentation and outreach to users.
- Data managements plans are becoming a requirement for NSF and DOE grant applications.
  - IT provided a general lab-wide data management plan.
  - The four halls provided hall specific plans based upon the IT plan.
  - All five plans are posted on the web for users to refer to.
    - links.

Computing requirements

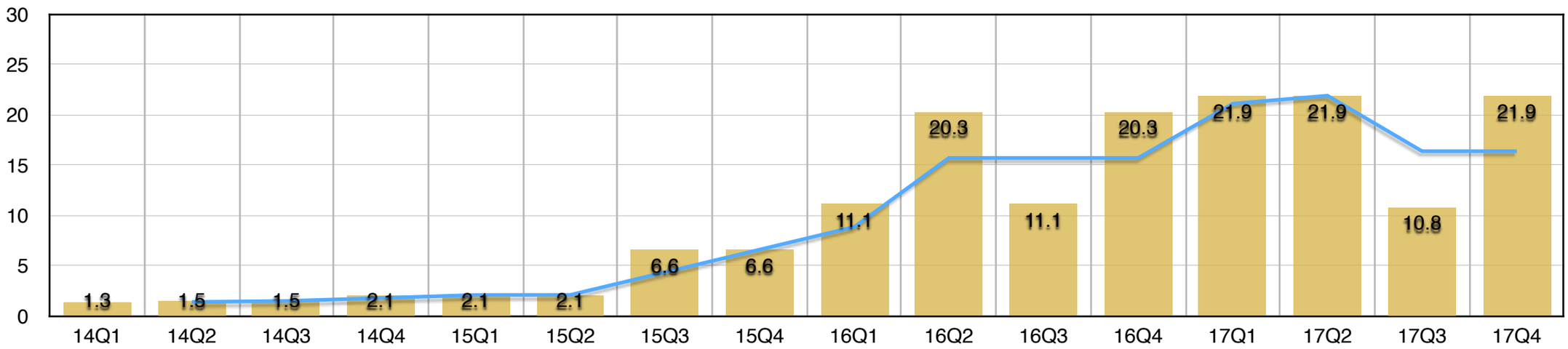
- The computing requirements for each hall are owned by the offline working groups and are driven by the scheduled experiments, the capabilities of the detectors, the data analysis workflow adopted by each hall and their analysis frameworks.
- Since the last review the Hall working groups have been working on benchmarks and data challenges that have allowed them to refine the input parameters to the computing requirements calculations.
- There has been a schedule re-baseline that has changed the timing of when computing resources are required, particularly hall-B.
- The tables on the next three slides present the CPU, disk and tape requirements broken down by quarter over the next four years.
- The throughput of the multi-threaded analysis frameworks developed by halls B and D seem to scale linearly with cores per node and clock speed. The standard unit measurement for the CPU requirement is a single compute core on one of the 2013 generation (2.0 GHz Sandy Bridge) compute nodes (16 cores per node).
- The assumption is steady state data taking so that in each quarter we must at least process one quarter's worth of data to not fall behind.

CPU

CPU in 1000 of 2013 equivalent cores

	2014				2015				2016				2017			
	14Q1	14Q2	14Q3	14Q4	15Q1	15Q2	15Q3	15Q4	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4
6 GeV	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	500	500	500	500	250	250	250	250
Hall A	8	8	8	8	14	14	14	14	17	17	17	17	17	17	17	17
Hall C	0	0	0	0	10	10	10	10	14	14	14	14	14	14	14	14
Hall B	189	377	377	566	566	566	566	566	566	9,772	566	9,772	11,613	11,613	566	11,613
Hall D	100	100	100	500	500	500	5,000	5,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000
Total/1000	1.3	1.5	1.5	2.1	2.1	2.1	6.6	6.6	11.1	20.3	11.1	20.3	21.9	21.9	10.8	21.9

CPU requirement (units of thousand 2013 equivalent cores)



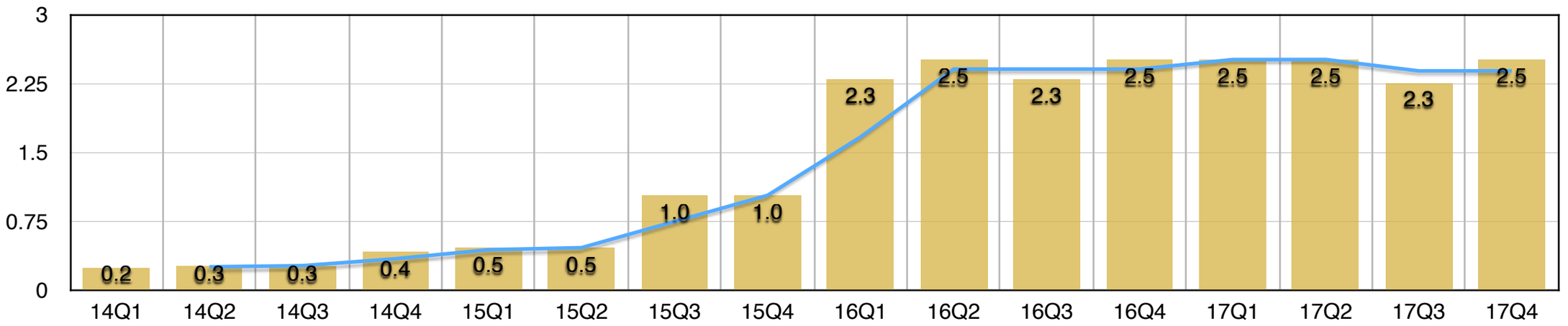
- Assuming performance scales with the number of CPU cores, how many cores of current vintage would need to be installed at the start of each quarter to handle the load in that quarter?
- There is still 6 GeV work so assume that starts with the current observed load and diminishes with time over the next four years.
- Assume large data challenges use resources borrowed from the LQCD clusters.
- For reference, the current cluster is ~1500 cores

Disk

Disk - volatile + work in TB (total in PB)

	2014				2015				2016				2017			
	14Q1	14Q2	14Q3	14Q4	15Q1	15Q2	15Q3	15Q4	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4
6 GeV	180	180	180	180	180	180	180	180	90	90	90	90	50	50	50	50
Hall A	12	12	12	12	30	30	30	30	127	127	127	127	127	127	127	127
Hall C	0	0	0	0	27	27	27	27	41	41	41	41	41	41	41	41
Hall B	26	51	51	77	77	77	77	77	77	284	77	284	325	325	77	325
Hall D	25	25	25	150	150	150	720	720	1970	1970	1970	1970	1970	1970	1970	1970
Total/1000	0.2	0.3	0.3	0.4	0.5	0.5	1.0	1.0	2.3	2.5	2.3	2.5	2.5	2.5	2.3	2.5

Disk requirement - work + volatile in PB



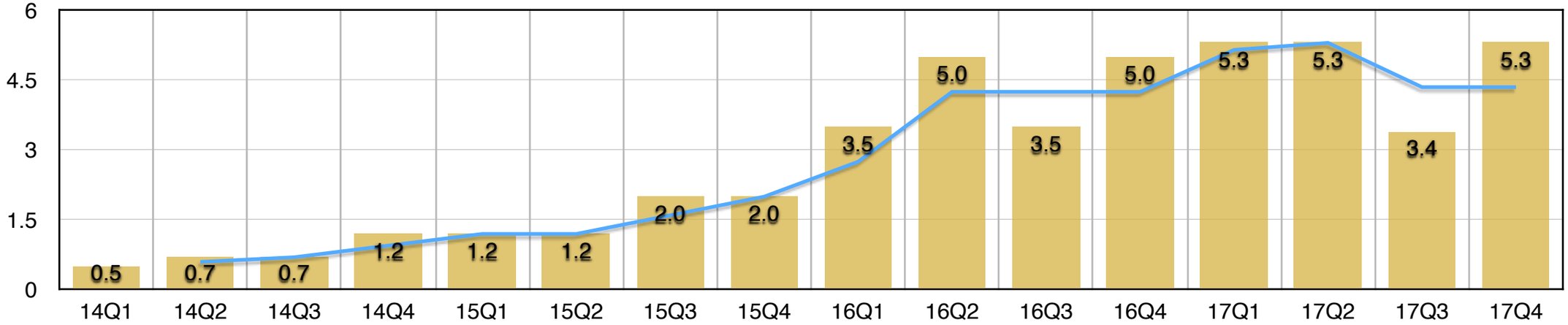
- Disk is in three flavors, work, volatile and cache.
  - Cache disk is part of the mass storage and isn't listed here.
  - Volatile and work have different cost per TB, see Sandy's talk.
  - The amount of both flavors of disk depends upon the volume of raw data and the analysis workflow.

Tape

Tape TB per quarter (total in PB/Q)

	2014				2015				2016				2017			
	14Q1	14Q2	14Q3	14Q4	15Q1	15Q2	15Q3	15Q4	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4
6 GeV	200	200	200	200	200	200	200	200	100	100	100	100	50	50	50	50
Hall A	20	20	20	20	60	60	60	60	250	250	250	250	250	250	250	250
Hall C	0	0	0	0	0	0	0	0	350	350	350	350	350	350	350	350
Hall B	253	505	505	758	758	758	758	758	758	2313	758	2313	2624	2624	758	2624
Hall D	0	0	0	200	200	200	1000	1000	2000	2000	2000	2000	2000	2000	2000	2000
Total in PB/Q	0.5	0.7	0.7	1.2	1.2	1.2	2.0	2.0	3.5	5.0	3.5	5.0	5.3	5.3	3.4	5.3

Tape requirement in PB/quarter



- Tape storage costs are controlled by three factors:
  - Cost of library infrastructure and maintenance.
  - Cost of “shelf space” in the library.
  - Cost of media.
- We store duplicate copy of the raw data outside the library.
- Media cost is paid from operating budget of hall.
- Must eject processed and raw data quickly to keep library costs down. - hand over to Sandy

Cost breakdown by quarter

Schedule of cpu and disk purchases.

	2014				2015				2016				2017			
	14Q1	14Q2	14Q3	14Q4	15Q1	15Q2	15Q3	15Q4	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4
CPU req k cores	1.3	1.5	1.5	2.1	2.1	2.1	6.6	6.6	11.1	20.3	11.1	20.3	21.9	21.9	10.8	21.9
Disk req PB	0.2	0.3	0.3	0.4	0.5	0.5	1.0	1.0	2.3	2.5	2.3	2.5	2.5	2.5	2.3	2.5
Tape req PB/Q	0.5	0.7	0.7	1.2	1.2	1.2	2.0	2.0	3.5	5.0	3.5	5.0	5.3	5.3	3.4	5.3
CPU $	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Disk $	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Tape $	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

- CPU and disk is procured from ENP Ops budget, tape media from hall Ops.
- Cost per unit of CPU, tape and disk is falling with time aim to purchase as late as possible.
- Avoid a large single year bump in spending by using the boundaries between fiscal years.

Conclusions

- There will be some concluding remarks here!
- The requirements provided by the halls seem reasonable and are consistent with what they have previously presented.
- The parameters used to calculate the requirements are constantly being refined.
- The schedule re-baseline has allowed us to push forward some computing procurements across fiscal year boundaries. Also to take advantage of decreasing unit costs.
- IT division are used to procuring and operating large clusters.
  - In 2014/15 the ENP cluster will be small compared to LQCD, can borrow resources for data challenges etc.
  - In 2016 and beyond ENP and LQCD will be approximately equal in size.
- Say something about the overall cost here but don't know what that will be yet…