Richard Gerber
NERSC Senior Science Advisor to the Director

April 29, 2014
Requirements Reviews

1½-day reviews with each Program Office
Computing and storage requirements for next 5 years

- Participants
  - DOE Program Managers & ADs
  - Leading NERSC users & key potential users
  - NERSC staff

Scientific Objectives
Computing, Storage, Software, Services Requirements
Reports From 9 Requirements Reviews Have Been Completed

- Computing and storage requirements for 2014 & 2017
- Executive Summary of requirements
- Case studies
- Second round, for 2017 requirements, will be completed in April 2014 (NP)

http://www.nersc.gov/science/hpc-requirements-reviews/reports/
Impact

• **Scientific justification for ASCR budget requests**
  - Quantitative requirements
  - Description of scientific benefit
  - Documented needs from science teams

• **Basis for NERSC 7 and NERSC 8 Mission Need**
  - Demand for hours beyond available resources
  - Science at scale, support for ensemble runs & HTC
  - Rich development environment

• **Guides NERSC services directions**
  - Application readiness teams
  - Queues for science at scale & HTC
  - Support for standard tools, libraries, applications
  - Planning for NERSC data services
Meeting Goals

• Gather computing, storage, and HPC services required to support NP research through 2017
• Collect a set of project-based “case studies” with scientific goals and how HPC requirements support achieving those goals
• Before we leave: high-level findings
• Ultimately: a written report for DOE
PRRs Influenced the Selection of Edison & N8

• Findings from first round of PRRs (Program Requirements Reviews)
  – The NERSC community would not be ready to effectively use accelerators in production by 2014
  – There is a need for improved I/O rates and disk storage
  – Many codes benefit from more memory per node, faster single-processor performance, and a high-bandwidth, low-latency interconnect
  – Productivity is more important than “feeds and speeds”

• PRRs findings formed basis of NERSC 7 Mission Need Statement
  – Edison has fast commodity Intel x86 processors, 64 GB/node memory, 6+ PB of /scratch, and novel high bandwidth, low-latency Aries interconnect
  – Adoption by NERSC community was immediate, with little porting effort
  – Performance is running 2X-4X that of Hopper on a per-core basis
Process

• Collect and refine requirements for 2017
  – Case study worksheets
  – Discussions at this meeting lead to high-level findings
  – Post-meeting refinement of case studies

• Draft a written report (Richard & Harvey)
  – Assemble case studies and check for internal consistency and compare against historical trends
  – Aggregate requirements and summarize
  – Create draft report for you & NP to review

• Send final draft to NP & ASCR for final approval

• Publish final report
Strategy Overview

Tell how these are needed to achieve your scientific goals – as specifically as possible

– Computational and storage resources
– HPC services
– Software

Additional important info: Are your codes ready for many-core? If not, what do you need?
Quantitative Method

• **Quantitative requirements are very important**
  – Hours needed
  – Archival data (HPSS) storage needed
  – Disk storage needed

• **For hours and archival storage**
  – Requirements from this review are summed
  – Scaled to full NP need by the fraction of 2013 NP usage represented by case studies
  – Important: Associate each case study with 2013 NERSC repo or repos
  – New/potential projects’ requirements added in separately

• **Like to do the same for Scratch and Project shared disk**
  – Please state 2013 usage and 2017 need so we can create a ratio
• The unit of “Hour” is defined as 1 Hopper core hour

• Please state your requirements in these units
  – How much computing will you need in multiples of a Hopper hour?
  – For this exercise, ignore the architecture – we will normalize this when future systems arrive, based on average application performance

• Give your best estimate for 2017 specifically
  – Remember that each year’s usage has historically been 2X the previous year’s
### Computational Hours

#### NERSC and NP Computational Hours Used

- **NERSC Trend**
- **NP Trend**
- **NP Usage**
- **All NERSC Usage**
- **NP Review Target 2014**
- **All Reviews Target 2014**

![Graph showing computational hours](image)

- **2002**
- **2003**
- **2004**
- **2005**
- **2006**
- **2007**
- **2008**
- **2009**
- **2010**
- **2011**
- **2012**
- **2013**
- **2014**
- **2015**
- **2016**
- **2017**

- 2.02 X/year
- 2.08 X/year
- 216 M used
- 4.9 B need
- 86 B trend
- 9.7 B trend
- 15.7 B need
- 2.1 B used
Data Storage Requirements

• **Archival storage estimate for 2017**
  – This is an aggregate number, not what you will add in 2017
  – Historical trend: $1.5-1.7X$ / year

• **Scratch (temporary)**
  – What is the maximum you will need at any given time during 2017?
  – Not just what you will need for a single run

• **Project shared disk space (permanent)**
  – What will you need for source code, data files or executables that will be constantly accessed and/or shared, etc.
Logistics: Schedule

- Agenda on workshop web page
  - http://www.nersc.gov/science/requirements/NP/
- Mid-morning / afternoon break, lunch
- Today: Case study presentations & discussions
- Self-organization for dinner
- Wednesday: overview, review, and discuss key findings
- Report: NP Intro + PI case studies + NERSC summary
  - Final Case Studies due June 1
  - Richard / Harvey review
  - PI/DOE draft review August 1
  - Final: September 1
- Final reports from previous workshops on web
  - http://www.nersc.gov/science/requirements
Logistics

• Get your presentations to us (Harvey/Richard)
  – Email
  – Web download
  – USB stick

• The laptop at the front will be used to display presentations

• We will stay on time
  – Descriptive and concise science justification
  – Please emphasize requirements and application readiness for manycore
  – Watch countdown timer
Questions?
Terms

- **“Memory”**
  - Volatile or “RAM”
  - Each “node” has a pool of RAM shared among all cores on the node
  - “Global memory requirement” means the sum of all the RAM on the nodes on which your job is running

- **“Many Core”**
  - “Processors” with 100s+ of “light-weight” cores
  - Slower clock speeds (energy efficient)
  - Not self-hosted; need a master CPU (today)
  - Special ways needed to write programs
  - GPUs and Intel Phi
Storage Terms

• “Scratch storage”
  – Temporary, purged after ~6 weeks
  – Fast: 10s – 100s of GB/sec
  – Not backed up
  – Access from a single system (at least at high performance)
  – Default quotas: ~ 10s TB + today

• “Permanent storage”
  – Not purged
  – Usually backed up (feasible into the future?)
  – Somewhat less performant
  – Maybe sharable
  – Center-wide access
  – Default quotas: ~10s GB (Home) to ~10-100 TB (Project) today

• “Archival Storage”
  – Permanent & long term
  – Much slower access time
  – No quotas: up to 10 PB today

Burst Buffers