ASCR Requirements Workshop

Initial Requirements Summary

Richard Gerber
Harvey Wasserman
NERSC User Services

January 6, 2011

U.S. DEPARTMENT OF ENERGY | Office of Science

NERSC National Energy Research Scientific Computing Center

BERKELEY LAB Lawrence Berkeley National Laboratory

Why Are A Clear Set of Requirements Important?

- So ASCR can evaluate what is needed to support its research mission.
- So NERSC can provide the necessary computers, storage, networking, software, and services.
- So you will have access to the HPC resources you need to perform your research.

Key Sections of Final Report

- **Executive Summary**
  - Brief Summary of Major Findings
- **Findings**
  - Summary of requirements
  - Top 4-5 overall requirements with sub-bullets
- **Case Studies**
  - Project / Code Description
  - HPC Requirements
    - Narrative
    - Summary Table
  - Support Services & Software

Initial Requirements Summary

- What we heard wrt requirements …
- Tell us what we heard wrong (and right!)
- Tell us what we missed

Data Analysis

- **Some in-situ analysis needs to be performed on big machine**
  - 1%?,10%?, 20%? increase in nodes / time for application codes??
  - Software development needed. How much allocation to support this? (ASCR allocation 2X of current % of NERSC?)
  - Size and number of debug runs will increase
  - Time for testing new app code version that includes in-situ analysis
- **A post-processing platform at NERSC is needed to be close to data**
  - Hardware requirements?
  - Software development needed. (Software infrastructure.)
  - Large SMP? (see later slides)
  - Need to move some data from remote sites to NERSC for analysis. Network requirements? ("No data islands"; ability to move large data between HPC sites)
- **Some data needs to be delivered to remote sites**
  - Network requirements? Work closely with Esnet to ensure the most advanced networking technologies are in place.
  - Workflow software requirements?

Data Analysis

- **Tools needed for analysis workflow (on supercomputers)**
  - Do they need to be developed? What is needed for this? Communities develop and support and NERSC integrate if needed.
  - Existing ones installed / maintained @ NERSC? NERSC help it run correctly and efficiently.
  - Workflow support experts needed to supply “workflow service”
- **Need to track simulations as they run**
  - Kepler support needed? (expert support needed)
- **Some data needs to be delivered to remote sites**
  - Network requirements?
  - Workflow software requirements?

Data

- **Sharing of simulation data is valuable**
  - NERSC should support integration of community portals into its environment (back-end PBS integration, authentication, data movement, metadata, etc. e.g. ESG)
  - Public (or authenticated) access to data repository at NERSC?
  - NERSC-provided portal to data? (yes! Collaborative portal development (limited). Deal with metadata and transfer.)
  - Add question to ERCAP: Do you want a "portal" to share your application data?
- **Need mathematicians and computer scientists to work together to develop data / analytics software**
  - What does NERSC need to provide or support?
- **Example: C/R optimization**
  - Need disk space 100X size of checkpoint files to generate performance statistics and verify methods
- **Expert data analysis as a service NERSC supplies**

I/O

- **NERSC consulting is valuable**
- **Support for I/O libraries is needed**
  - HDF5, netCDF, parallel netCDF, ROMIO, MPIIO, ADIOS?
  - Access to low-level functionality is valuable: Lustre striping, fs APIs?, documentation
- **Need to test at full scale (application scale)**
  - How much allocation needed for this? Hard to predict? (see R.Ross email)
  - Dedicated time for testing (iterative testing)
- **Need mechanism for getting middleware improvements into production**
  - What needed? (Proactively work with vendors)

Visualization

- **Analysis usually (often?) needs same supercomputer as simulation**
  - Why? (quantitative argument?)
  - Cluster analysis software for this doesn't exist today?
- **Need access to supercomputers and vis machines**
  - Architectural requirements for vis machine? (carver/magellan-like OK)
  - Vis machine needs local access to data: shared /scratch with SCs
- **Remote visualization**
  - High-speed network needed (how high?)
  - Distributed storage (replicated data (chunks) managed intelligently. existing? Needs to be developed? Common data repository with individual files geographically distributed. - arie)
- **Need large SMP to test algorithms and codes before implementing distributed-memory version**
  - SMP memory needed?

Vis / Analysis

- **Need big machines and emerging platforms to try out algorithms**
  - Allocation needed?
- **I/O speed needs to keep pace with size of machines**
  - Can this be quantified? (e.g. I/O bw vs. system memory size)
- **Portals (etc)**
  - Web servers, email lists, wikis
- **NERSC needed as a vehicle to deploy software**
  - Support for installing and supporting software (by third parties?)
- **Access to testbeds: GPU clusters (hybrid), cloud testbeds, FPGAs (wes)**

John Bell

- **Async (farmed-out) in-situ data analysis needed (can't write all data to disk)**
  - Extra allocation needed for development? Then in production?
- **Need a better programming model (to express data locality)**
  - What requirement(s) can this be translated into?
  - Provide access to PGAS on future architectures. Provide training.
  - Assure platforms support development of UPC, etc. (documentation of, access to, low-level functions)
- **Disseminate data to broader community**
  - NERSC portal?
- **Can integrate intrusive UQ with 50X resources**
- **Architectural requirements listed already (8 GB/node, 30X increase in hours, e.g.)**

Larsson / Lele

- **NERSC important for developing code and scaling up to largest systems**
  - Large-scale reimbursement
- **Need to do runs at all scales**
  - What requirement(s) can this be translated into?
- **More hours needed to include chemistry**
  - Would be game-changer
  - Memory requirements to hold tables
- **Post-processing is a bottleneck due to queue waits**
- **Need common filesystem so don't have to move data to analysis machine**
- **Need reliable systems (no long queue waits, then have job fail)**
- **Need help choosing programming models; can't waste limited development resources**
- **NERSC portal to distribute data**
- **Help with visualization**

UQ

- **UQ becoming more important (?)**
  - E.g. climate, combustion, subsurface flow, accelerator design
- **Job management needed for ensemble runs**
  - Intelligent scheduler, monitoring, fault management
  - New software needed? Express this as a requirement?
- **UQ is a project unto itself**
  - What allocation / resources needed?
  - “Can integrate UQ with 50X resources”
  - UQ takes 6mo – 1 yr typically (limited by workflow?)
- **UQ as a driver for implementing workflow tools into HPC**

Performance Evaluation / CS Research

- **Exclusive access to many different architectures needed**
  - CS: need access to smaller system
  - Bigger runs for performance eval
    - Need support from center to enable runs at full scale
- **More large-scale tests needed to study issues at 10X+ concurrency**
- **Autotuning for multi-node will require large allocation**
  - Large search space, even for single-node on modern complex architectures
- **Detailed documentation for low-level stack functions for CS research**
- **Easy access to reliable power readings at node level**
- **Fast turnaround for small jobs**

Esmond Ng

- Graph partitioning software
- PETSc, ScalaPACK, etc.
- Need large HPC allocation when doing scaling studies
- Large memory nodes (12GB?)
- Fast turnaround for development and testing
- Help with programming models, using hybrid systems,…

Lawrence Berkeley
National Laboratory

Programming Models Research

- Access to low-level APIs
- Access to development machines (running production software) that can be crashed
- Make new programming models available to application developers
- Virtualization for development and testing in different enviroments
- (All applicable to OS research also)

National Energy Research Scientific Computing Center