Data Requirements from NERSC Requirements Reviews

Richard Gerber and Katherine Yelick
Lawrence Berkeley National Laboratory

Summary

Department of Energy Scientists represented by the NERSC user community have growing requirements for data storage, I/O bandwidth, networking bandwidth, and data software and services. Over the next five years, these requirements are well above what would be provided by increases that follow historical trends. This report focuses primarily on the data needs of the modeling and simulation community, which dominates NERSC usage, although NERSC also has projects that involve data analysis of observational data sets. Regardless of the source of data, scientists across all fields recognized that qualitatively new challenges are arising in storing, accessing, sharing, managing, and analyzing massive data sets. Researchers also point to the need for new data analysis algorithms, networking capabilities, and the ability to use emerging hardware and software platforms. Data analytics are subject to the same architecture challenges plaguing the rest of computing, namely, the lack of clock speed scaling and growing memory and I/O imbalance, but with algorithms that may require an even higher need for high data rates and random access that physical simulations.

The growth in demand for data systems and services is coming from multiple sources:

- The increased size, resolution, and complexity of large-scale simulations are producing 100s of terabytes of data per simulation. In some cases, in-situ analysis is being developed to analyze data as the simulation runs.
- Data is being imported into NERSC from experimental facilities and remote observations with data rates growing faster than Moore's Law.
- A growing appreciation for the intrinsic value of scientific data, both simulated and observed, has increased interest in serving scientific data sets to the broader community. This improves the reproducibility of science, enables multiple independent discoveries from a single data set, and allows for analysis across data sets, e.g., combining observed and simulated data.
- Massive ensemble simulations are increasingly used to quantify uncertainty, to screen a related set of materials (The Materials Genome), proteins, or other domains require new techniques for storing, indexing, search

These challenges touch on what is commonly referred to as the "V's"[1] of data-intensive computations:

---

[1] Some lists combine *variability* and *veracity* or omit *value* to produce a list of 4 V's.

- Volume: petabyte data sets projected to exceed normal growth plans;
- Velocity: the speed with which data arrives, which leads to on-demand data processing for experiments and in-situ analysis for simulations;
- Variability: the ability to combine distinct data sets together, e.g., simulation and experimental data or data from different types of simulations;
- Veracity: data that is noisy or incomplete, which arise from inaccuracies in data measurement (e.g., short-read sequencers, missing CMB background behind the Milky Way), floating point approximations or missing features in a computational model; and
- Value: irreplaceable observational data and community data sets that require particular expertise, hardware, and software configurations that make it impractical for an individual scientist to reproduce.

Background

In a series of six High Performance Computing and Storage Requirements Reviews between 2009-2011 NERSC, DOE Program Managers, and leading domain scientists derived a set of production computing needs for the Office of Science. The results are available in a set of reports at http://www.nersc.gov/science/hpc-requirements-reviews/. This first round of reviews targeted needs for 2014. A second round is underway, gathering requirements for 2017. Second round reviews with the offices of Biological and Environmental Research (BER) and High Energy Physics (HEP) were completed in 2012 and the reports are in preparation. Reviews with the remaining four offices will be completed by early 2014.

The first round of reviews revealed a need for 15.6 billion hours of computing time in 2014, more than 10 times greater than can be provided by Hopper, the current production supercomputer for the Office of Science. Data requirements from the reviews for 2014 were a secondary consideration, primarily because the reviews concentrated on users with the largest computational demands, most of them from NERSC's traditional simulation community. Nonetheless, scientists in every review expressed their concerns regarding data: they fully and explicitly recognized that data issues – storage volume, I/O rates, data management, and data analytics – were quickly growing beyond their ability to deal with them effectively. Data, which was once an afterthought to many, was becoming a factor in their simulation workflows that they could no longer ignore.

In addition to the data problems derived from the modeling and simulation workload, NERSC also has substantial projects that involve experimental or observational data sets, including those from the Joint Genome Institute, the Large Hadron Collider at CERN (Alice and ATLAS), the Palomar Transient Factory, the Planck project for measuring Cosmic Microwave Background (CMB), the 20th Century Climate Re-analysis Project, the Daya Bay Neutrino Experiment (housed in China), the STAR experiment at BNL/RHIC. These projects use hardware, staff and services that use NERSC's data infrastructure including the NGF high performance parallel filesystem, HPSS archival tape storage, optimized data transfer nodes, and Science Gateway services for convenient external data access. These projects are not entirely supported by the NERSC Facility budget from ASCR, but instead involved direct support from other program offices (BER, HEP, NP, NASA) for hardware, software, and staff support.

In the second round of requirements reviews, NERSC is including representatives from data-intensive projects and also gathering input from new communities that have extreme unmet data requirements. The goal is to make data forecasts as reliable as the projected needs for computing cycles.

Requirements

Given the caveats, it is still possible to use results from the requirements reviews and reasonable extrapolations to arrive at both quantitative and qualitative production data needs for the Office of Science. These requirements largely reflect the needs of the simulation community, with some influence from data intensive projects that are currently using NERSC. They do not include the needs of communities that currently do not have a presence at NERSC, even though these groups, e.g. BES light sources and HEP users of accelerators and telescopes, likely have unmet data requirements that outstrip even those from existing NERSC users.

For reference, as of late 2012, NERSC provides about 10 PB of "spinning disk" and about 24 PB of archival storage for its users, with an archive capacity of over 40 PB.

Qualitative Needs

Attendees at the requirements reviews expressed needs for data storage, I/O bandwidth, and data management tools far beyond today's capabilities. Major concerns included having enough "fast access" data storage (not tape), I/O bandwidth to support checkpointing and simulation output (I/O should not exceed approximately 10% of a simulation's total run time), the current lack of data management tools, and a growing need for hardware and software to support data analytics.

Data needs were prominent in the high-level summary of findings and requirements in each report from the requirements reviews. Below is a brief summary from each program office.

BER (target 2017 draft): "The key needs are … access to more computational and storage resources … and the ability to access, read, and write data at a rate far beyond that available today …"

HEP (2017 draft): "The key needs are … more computing cycles and **fast-access** storage; support for data-intensive science, including: improvements to archival storage, analytics (parallel, DBs, services, gateways etc.), data sharing, curation, and provenance …" Slow data access for data stored on tape is a major concerns for many HEP science team; "fast-access" data storage was deemed to be a priority.

ASCR (2014): "Applications will need to be able to read, write, and store 100s of terabytes of data for each simulation run. Many petabytes of long-term storage will be required to store and share data with the scientific community."

BER (2014): Many BER projects have mission-critical time constraints; examples include predictions for the next IPCC climate change report and the Joint Genome Institute's need to maintain a four-month update cycle for genome datasets. Such projects demand a computational infrastructure that includes powerful, yet highly reliable, resources… and resource reservation policies.

BER (2014): Data manipulation and analysis is itself becoming a problem that can be addressed only by large HPC systems. Simulation output will become too large to move to "home" institutions; therefore, NERSC needs to integrate robust workflow, data portal, and database technology into its computational environment and significantly increase real-time-accessible data storage capacity.

BES (2014): " [There is a need to support] … huge volumes of data from the ramp-up of the SLAC LINAC Coherent Light Source (LCLS) [and other experimental facilities in BES]."

HEP (2014): Science teams need to be able to read, write, transfer, store online, archive, analyze, and share huge volumes of data. 1. The projects considered here collectively estimate needing a 10-fold increase in online disk storage space in three to five years. 2. HEP researchers need efficient, portable libraries for performing parallel I/O. Parallel HDF5 and netCDF are commonly used and must be supported. 3. Project teams need well-supported, configurable, scriptable, parallel data analysis and visualization tools. 4. Researchers require robust workflow tools to manage data sets that will consist of hundreds of TB. Science teams need tools for sharing large data sets among geographically distributed collaborators. 5. The NERSC Global File System currently enables high performance data sharing and HEP scientists request that it be expanded in both size and performance. 6. Scientists need to run data analysis and visualization software that often requires a large, shared global memory address space of 64-128 GB or more. 7. Researchers anticipate needing support for parallel databases and access to databases from large parallel jobs.

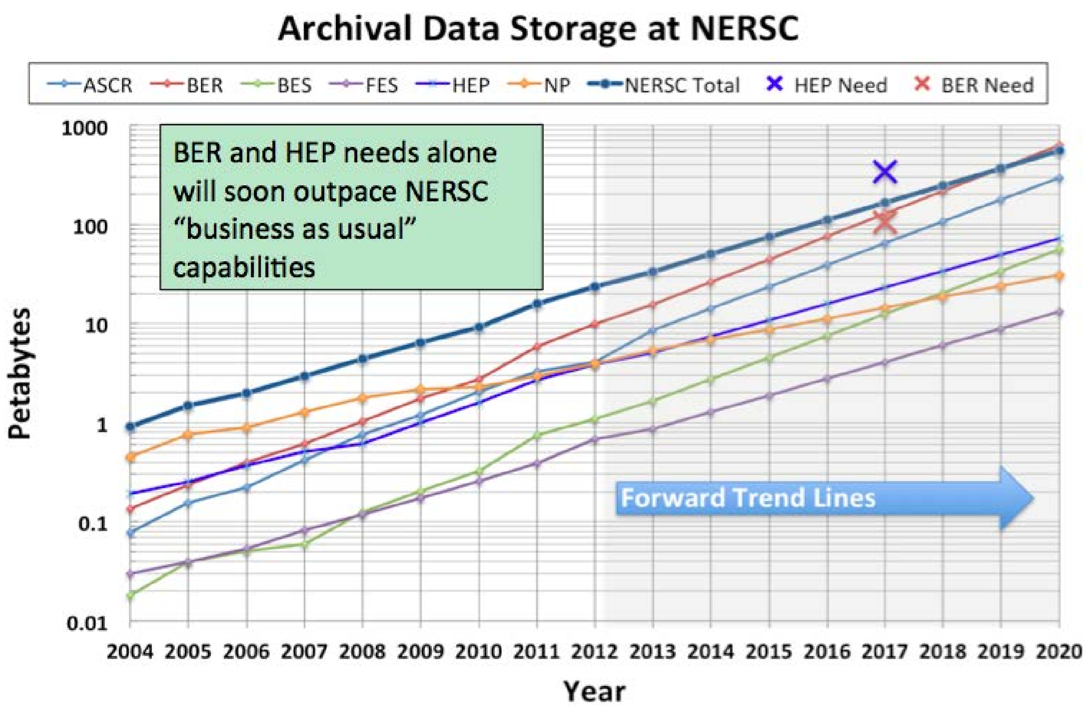
FES (2014): "[Researchers need] data storage systems that can support high-volume/high-throughput I/O."

NP (2014): "[Needs include] useable methods for cross-correlating across large databases … grid infrastructure, including the Open Science Grid (OSG) interface. Increased [computing] capacity … has resulted in … a significant increase in I/O demands in both intermediate and long-term storage. "

Quantitative Requirements

We chose to focus on archival storage, data that is deemed to be of permanent value and will be saved indefinitely. This is will serve as a proxy for all data, both archival and "live." Archival data at NERSC is currently stored in an HPSS system this is fronted by a large disk cache, but ultimately relies on tape storage for the long term. We can derive quantitative numbers for archival storage because NERSC has detailed historical data on HPSS usage and most attendees at the requirements reviews gave projections for their archival storage needs.

The figure below shows historical HPSS usage through 2012 and linear projections for each office (calculated independently) through 2020. Tentative estimates from the second round of requirements reviews from BER and HEP are also shown. This plot only represents "user" data, omitting "spinning disk" backup data (currently more than 12 PB), as well as data stored by NERSC as "overhead" (several PBs).



Archival data storage has consistently grown by a factor of 1.7 year over year for over a decade. Tentative results of the second of reviews indicate that this will not be adequate moving forward. Needs from just HEP and BER will be much greater than what would be provided by accommodating the historical trend. In BER the demand comes largely from climate and genomics research while the need in HEP is driven by simulated data required to support and interpret measurements from accelerators, telescopes, and satellite missions.

The following table gives the amount of archival data currently stored at NERSC for each office and projections from the workshops, given all the caveats related above. The 2014 numbers are very rough estimates; the projections for 2017 from the second round of workshops are expected to more faithfully represent all Office of Science production data needs.

Table 1 Archival Storage on NERSC's User HPSS System

Office	2012 Usage (PB)	2014 Projected Need (PB)	2017 Projected Need (PB)
ASCR	4.1	7.7	
BER	10.1		106*
BES	1.1	9.7	
FES	0.7	1.8	
HEP	3.9		340*
NP	4.0	28.6	

*Preliminary results from 2012 requirements reviews.

Research Problems in Data-Intensive Science (ASCR Report)

There is more detailed information about data needs of particular science projects described in the case studies in each of the six phase 1 reports. Here is (verbatim) an overview of data analytics and needs from the ASCR report, as it is particularly relevant and touches on some of the overarching data challenges across science disciplines and the associated research problems.

With the anticipated improvements in both experimental and computational capabilities, the amount of data (from observations, experiments, and simulations) will be unprecedented.

For example, by 2014 fusion simulations will use 1 billion cells and 1 trillion particles. Based on mean-time-between-failure concerns when running on a million cores, these codes will need to output 2 GBs/sec per core or 2 PB/sec of checkpoint data every 10 minutes. This amounts to an unprecedented input/output rate of 3.5 terabytes/second. The data questions to consider at the extreme scale fall into two main categories: data generated and collected during the production phase, and data that need to be accessed during the analysis phase.

Another example is from climate modeling where, based on current growth rates, data sets will be hundreds of exabytes by 2020. To provide the international climate community with convenient access to data and to maximize scientific productivity, data will need to be replicated and cached at multiple locations around the globe. These examples illustrate the urgent need to refine and develop methods and technologies to move, store, and understand data.

The data issue is cuts across all fields of science and all DOE Office of Science Program Offices. Currently, each research program has its own data-related portfolio; ASCR program managers envision an integrated data analytics and management program that will bring multi-disciplinary solutions to many of the issues encountered in dealing with scientific data.

In Applied Mathematics Research, data analytic needs include

- Improved methods for data and dimension reduction to extract pertinent subsets, features of interest, or low-dimensional patterns, from large raw data sets;
- Better understanding of uncertainty, especially in messy and incomplete data sets; and
- The ability to identify, in real time, anomalies in streaming and evolving data is needed in order to detect and respond to phenomena that are either short-lived or urgent.

In Computer Science Research, issues being examined include:

- Extreme-scale data storage and access systems for scientific computing that minimize the need for scientists to have detailed knowledge of system hardware and operating systems;
- Scalable data triage, summarization, and analysis methods and tools for in-situ data reduction and/or analysis of massive multivariate data sets;
- Semantic integration of heterogeneous scientific data sets;
- Data mining, automated machine reasoning, and knowledge representation methods and tools that support automated analysis and integration of large scientific data sets, especially those that include tensor flow fields; and
- Multi-user visual analysis of extreme-scale scientific data, including methods and tools for interactive visual steering of computational processes.

Next-generation Networking Research is concerned with

- Deploying high-speed networks for effective and easy data transport;
- Developing real-time network monitoring tools to maximize throughput; and
- Managing collections of extreme scale data across a distributed network