More Data, More Science... and Moore’s Law?

Kathy Yelick

Associate Laboratory Director for Computing Sciences
Lawrence Berkeley National Laboratory
Professor of Electrical Engineering and Computer Sciences
University of California at Berkeley
Science is poised for transformation
The Legacy of Team Science

Radiation Lab staff on the magnet yoke for the 60-in cyclotron, 1939, including:
E. O. Lawrence
Edwin McMillan
Luis Alvarez
J. Robert Oppenheimer
Robert R. Wilson
17-year-old Brittany Wegner creates breast cancer detection tool that is 99% accurate on a minimally invasive, previously inaccurate test.

Machine Learning + Online Data + Cloud Computing
Experimental Science is Changing

JAX® MICE & SERVICES

JAX® Mice are the highest quality and most-published mouse models in the world. Take advantage of our large inventories of common inbred strains and the convenience of having your breeding and drug efficacy needs met by the leading experts in mouse modeling.
NERSC: State-of-the art supercomputing for the broad science community – over 7000 users, 700 applications mostly in simulation
Computers are used to understand things that are for experiments alone, so simulations are used.
“Big Data” Changes Everything…What about Science?
About 10,000 visiting scientists (~2/3 from universities) use Berkeley Lab research facilities each year, which provide some of the world’s most advanced capabilities in materials science, biological research, computation and networking.
Data Growth is Outpacing Computing Growth

Graph based on average growth
Old School Scientific Workflow
Computing, experiments, networking and expertise in a “Superfacility” for Science

Science at the boundary of theory and experiment ... simulation and data analytics
Integration of Simulation and Observational Science

Simulations aid in interpreting data


Intermediate Palomar Transient Factory with DESI, CMB-S4 and LSST coming

Image subtraction, machine learning in minutes
Re-Use and Re-Analyze Previously Collected Data

- **Materials Genome Initiative**
  - Materials Project: Over 10,000 users!
  - “World-Changing Idea of 2013”

Computers programs run by “bots”
Real-Time Analytics in Health

3 min goal (1 sec/iteration)

Michael Driscoll HPC optimization

Compressed Sensing Approach by Mike Lustig et al
MRI results Wenwen Jiang
Data and Simulation in the environment

New climate modeling methods, including AMR "Dycore" produce new understanding of ice

Understand interactions between environmental microbiomes and climate change with *kilometer resolution models* that track dynamic 3D features (with AMR) and *genome-enabled analysis* of environmental sensors.
Figure 7: Hourly averaged actual usage is shown on the left. And hourly averaged predicted usage is shown on the right. Triangles markers show the averaged temperature. As presented in Tables 4 and 5, the predicted usage shows higher values than the actual usage, demonstrating that differing pricing policies affect household usage patterns.

accurate short-term forecasts, our baseline model aims to capture intraday characteristics that persists for years. Our tests show that one of the boosting technique, GTB, could incorporate important features such as outdoor temperature and capture the core user behavior. For example, the baseline model from GTB accurately reproduces the lag between the daily peak temperature and peak electricity usage.

The ultimate objective of our work is to evaluate the effectiveness of the different pricing schemes. The new baseline is an important component. This preliminary work demonstrate that new approach is promising, but additional work is needed to evaluate the effectiveness of this approach. For example, we should to re-evaluate the features used in the regression models and systematically measure their impact.

Acknowledgment

This work is supported in part by the Director, Office of Laboratory Policy and Infrastructure Management of the U.S. Department of Energy under contract No. DE-AC02-05CH11231. This work is also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) grant funded
Computing and network facilities need to adapt
Superfacility: Integrated network of experimental and computational facilities and expertise

Computing and Data Facilities

Experimental Facilities

A single interconnected “facility” where data is acquired, stored, analyzed and served

ESnet

User Community

Light Sources

Sequencers

Telescopes

Particle Detectors

Microscopes

Expertise

Methods, models, analytics, and software

Execution plan: one science area at a time

Applied Math

VISIT
ESnet: Data driven science drives network capacity

Science DMZ to deliver bandwidth to the end users
OSCARS for bandwidth reservation

100 Exabytes/year by 2024!
Systems configured for data-intensive science

NERSC Cori has data partition (Haswell) and pre-exascale (KNL) NVRAM file system with close to 2 PB at 2 TB/sec
WAN-to-Cori optimized for streaming data: 100x faster from LCLS to Cori and Globus to CERN
Real-time queue prototyped at NERSC

- In 1998 dedicated hardware; now prototype queue on Cori
- <1% of NERSC allocation
- Cryo-Em, Mass spec, Telescopes, Accelerator, Light sources

Cryo-EM: Image classification
Nogales Lab

PTF: Image subtraction pipeline

ALS: 3D Reconstruction, rendered on SPOT web portal
Interactive Analytics using Jupyter

Science notebooks through Jupyter (iPython)

- Widely used in science
- Interactive HPC LDRD

Deployed at NERSC:
- >100 users pre-production

Fernando Perez et al
Containers for HPC Systems

- Data analysis pipelines are often large, complex software stacks
- NERSC Shifter (with Cray), supports containers for HPC systems
- Used in HEP and NP projects (ATLAS, ALICE, STAR, LSST, DESI)
Old School Scientific Data Search

Tip: Try entering a descriptive word in the search box.

Image size: 153 × 133

No other sizes of this image found.

Visually similar images - Report images
Automated Search, Meta-Data Analysis, and On-Demand Simulation

Challenge 2: Search needs to account for scale and lineage of data and the I/O challenges of future systems.

Data search capabilities need to address scalability at various levels: a) machine learning algorithms must generate metadata at the rate and scale of the data volumes being generated; b) the metadata generation process must address the I/O challenges of the future exascale systems, and c) the metadata storage layer needs to address scalability.

Challenge 3: The complexities and intricacies of scientific data, as well as, machine and deep learning algorithms require a careful consideration of the human factors. Machine learning techniques can help with learning about the data and generating metadata. However, this is not sufficient for scientific data, since the complexity of the data often requires specialized domain knowledge and understanding. Automated metadata generated from machine learning algorithms will likely need to be curated by humans to ensure accuracy. Additionally, the machine learning model needs to understand the terms or signals that might arise from a user's query. Thus, it is important to understand how people interact and want to interact with scientific data search and machine-generated metadata labels.

Project Objectives: Designing a data integration ecosystem.

Our proposed techniques bring together a unique blend of skills that includes machine learning, human-computer interaction, and experience with scientific domains and users at facilities. Our goal is to make data a first-class discoverable resource at supercomputing centers through the powerful concept of search.

Figure 1 shows the conceptual system architecture that will be enabled by the research proposed in this proposal. The ScienceSearch framework has three key components: a) metadata generation, b) the Ground [38] metadata storage framework, and c) an interface layer. The metadata generation framework uses a variety of machine learning techniques to generate the context of the data from both application data, as well as system level information. Ground is a data context service that provides the metadata storage layer. The interface layer allows the users to interact with the system to verify and validate automated metadata generated.

We envision the ScienceSearch framework will be available at supercomputing centers and users can make their data available to the system. The ScienceSearch framework will use the data sets and, ecosystem artifacts associated with the data (e.g., proposals, workflow and system logs, publications) to learn and generate metadata labels. The ScienceSearch framework will use active learning to surface the metadata labels to users for feedback. The users can validate, add, delete or edit labels. Similarly, we anticipate that...

Automated metadata extraction using machine learning
Computational research challenges are substantial
Software implementations at scale in pipeline:

- **MicroCT imaging**
- **Segmentation**
- **Topological Analysis**
- **Analysis**
- **Simulation**
- **Visualization**
Automated detection and analysis of particle beams in laser-plasma accelerator simulations.

Fig. 5. Comparison of particle selection with/without MVEE: extracting the orientation and the axes of an enclosing ellipse from (a) produces (b), increasing the number of particles from 173 to 263. Colors indicate the density of particles, using only $(x, y)$-coordinates, and black dots show potential particles to belong to the beam, according to the different methods.

Maximum (beam candidate region) per time step. In addition, this is a way of accruing more samples and detecting secondary beams when these are almost as prominent as the primary beam, associated to the maximum of $f$.

During the searching for values that are approximately equal to $\max(f)$, we keep not only the maximum, but all bins where $f \geq u \times \max(f)$, where $u$ is an uncertainty or tolerance parameter, here empirically set to 0.85. While this value enables the detection of the main and the secondary beams (when present), lower values of $u$ could be used to control the amount of particles to be selected at a lower accuracy of beam position. From this point, we refer to the subset of particles conditioned to $u \times \max(f)$ and its adjacency, calculated for each time step, as "beam candidates".

Figure 4 (top) presents projections of Figure 3.b with their calculated beam candidates emphasized in red. These are the result of our first attempt to improve particle selection by using an algorithm known as minimum volume enclosing ellipsoid as in Khachiyan & Todd (1993), which is able to enclose previously selected particles and to include others based on a geometrically defined polytope. Figure 5 illustrates the algorithm when applied to LWFA data, showing the selected particles as black dots; these particles are not in the most dense region (red) once the colors refer to $(x, y)$-density calculation. When including compactness in $\text{px}$, the most dense region happens further ahead. As distinct from calculating center of mass and forcing an ad hoc diameter or semi-major/minor axes, the minimum volume enclosing ellipsoid (MVEE) algorithm [Khachiyan & Todd (1993); Kumar & Yildirim (2005); Moshtagh (2009)] takes the subset of points and prescribes a polytope model to extrapolate a preliminary sub-selection to other particles likely to be in the bunch. The MVEE algorithm is a semidefinite programming problem and consists of a better approximation to the convexity of subsets of...
<table>
<thead>
<tr>
<th>7 Giants of Data</th>
<th>7 Dwarfs of Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic statistics</td>
<td>Monte Carlo methods</td>
</tr>
<tr>
<td>Generalized N-Body</td>
<td>Particle methods</td>
</tr>
<tr>
<td><strong>Graph-theory</strong></td>
<td>Unstructured meshes</td>
</tr>
<tr>
<td>Linear algebra</td>
<td>Dense Linear Algebra</td>
</tr>
<tr>
<td>Optimizations</td>
<td>Sparse Linear Algebra</td>
</tr>
<tr>
<td>Integrations</td>
<td>Spectral methods</td>
</tr>
<tr>
<td>Alignment</td>
<td>Structured Meshes</td>
</tr>
</tbody>
</table>
Random Access Analytics

- Genome assembly "needs shared memory"

Global Address Space
- Low overhead communication
- Remote atomics
- Partitions for any structure

Scales to 15K+ cores
Under 10 minutes for human
First ever solution

Productive Programming

Spark

Speed
Run programs up to 100x faster than Hadoop MapReduce in memory, or 10x faster on disk.

- High failure rate
- Slow network
- Fast (local) disk

And Spark is still 10x+ slower than MPI
SPARK Analytics on HPC

Weak scaling (fixed problem size per node): Power Iteration Benchmark

SPARK on HPC vs. clusters

- Network, I/O, and virtualization all key to performance
- Increased scale from $O(100)$ to $O(10,000)$ cores

Chaimov, Malony, Iancu, Ibrahim, Canon, Srinivasan
Filtering, De-Noise and Compressing Data

AmeriFlux & FLUXNET: 750 users access carbon sensor data from 960 carbon flux data years

Arno Penzias and Robert Wilson discover Cosmic Microwave Background in 1965
How will we get enough computing for these problems?
Architectures for Data vs. Simulation

Different architectures for simulation? Can simulation use data architectures?
More Parallelism at Lower Levels

Application Performance Growth
(Gordon Bell Prizes)

Specialization
“Killer cellphones”

Manycore Parallelism

Rest of the world gets parallelism

Vector parallelism
“Killer micros”

1.0E+18
1.0E+17
1.0E+16
1.0E+15
1.0E+14
1.0E+13
1.0E+12
1.0E+11
1.0E+10
1.0E+9
1.0E+8
End of Transistor Density Scaling

ITRS now sets the end of transistor shrinking to the year 2021
Device alternatives require lower clock \(\rightarrow\) more parallelism

Today's CMOS Technology

Tunneling FET advantage *only at low clock rates*
Specialization: End Game for Moore’s Law

Google designs its own Tensor Processing Unit (TPU)

Intel buys deep learning startup, Nervana

NVIDIA builds deep learning appliance with P100 Tesla’s

FPGAs
Data processing with special purpose hardware

- General trend towards specialization for performance
- Data processing (on raw data) will be first in DOE

Particle Tracking with Neuromorphic chips
Computing in Detectors

Deep learning processors for image analysis

FPGAS for genome analysis
KATHY YELICK'S
2031:
a science odyssey
Life of a Scientist in 2031

- No personal/departmental computers
- Users don’t login to HPC Facilities
- Travel replaced by telepresence
- Lecturers teach millions of students
- Theorems proven by online communities
- Laboratory work is outsourced
- Experimental facilities are used remotely
- All scientific data is (eventually) open
- Big science and team science democratized
The scientific process is poised to undergo a radical transformation based on the ability to access, analyze, simulate and combine large and complex data sets.