More Data, More Science and…
Moore’s Law

NERSC 40th

Kathy Yelick
Associate Laboratory Director
for Computing Sciences
Lawrence Berkeley National Laboratory

David Skinner
NERSC Strategic Partnerships Lead
NERSC Strategy: Science First

- Response to scientific needs
  - Requirements setting activities
- Support computational science:
  - Provide effective machines that support fast algorithms
  - Deploy with flexible software
  - Help users with expert services
- NERSC future priorities are driven by science:
  - Increase application capability: “usable Exascale”
  - Simulation and data analysis
HPC Resources for User Science

DOE Big Data  *Volume, velocity, variety, and veracity*

**Biology**
- **Volume**: Petabytes now; computation-limited
- **Variety**: multi-modal analysis on bioimages

**High Energy Physics**
- **Volume**: 3-5x in 5 years
- **Velocity**: real-time filtering adapts to intended observation

**Cosmology & Astronomy**
- **Volume**: 1000x increase every 15 years
- **Variety**: combine data sources for accuracy

**Light Sources**
- **Velocity**: CCDs outpacing Moore’s Law
- **Veracity**: noisy data for 3D reconstruction

**Materials**
- **Variety**: multiple models and experimental data
- **Veracity**: quality and resolution of simulations

**Climate**
- **Volume**: Hundreds of exabytes by 2020
- **Veracity**: Reanalysis of 100-year-old sparse data
Top 15 Science Data Projects in NERSC Filesystem

Daya Bay
Urban Sensor + Sim
Supernova (PTF)
Cosmology Sim
Planck (CMB)
Climate 100
ALS (Light Source)
Climate Reanalysis
BAO
Alice (LHC)
SN Factory
STAR Detector
Extreme Weather
Materials Project
JGI (Genomes)

Biggest online data sets are from:

- Experimental facilities
- Observations
- Simulations
- Reconstructed observation
- Sensors

Total for these projects:
1.5 Petabytes of Disk
4.5 Petabytes of Tape
Data Growth is Outpacing Computing Growth

Graph based on average growth
NERSC and Esnet: WAN data trends

Daily WAN traffic in/out of NERSC over the last decade

Roughly 10x 2011-2016

Automated data pipelines for large scale genomics, LHC, image processing

Community access to data and analysis, gateways

Data at NERSC is secure, reliable, fast, open, flexible
Data from DOE facilities: Tomorrow is already here

Experimental facilities will be transformed by high-resolution detectors, advanced data analysis techniques, robotics, software automation, and programmable networks.

1. Detectors capable of generating terabit data streams.
2. Data reduction & feature extraction *in situ*, using advanced algorithms and programmable hardware.
3. Increase scientific throughput from robotics and automation software.
4. Computational tools for analysis, inter-comparison, simulation, visualization.
5. Data management and sharing, with granular access control.
Transforming Science: Finding Data
(Just a Few) New Data Methods: Tools vs. APIs

Simulation, data analysis, and visualization tools integrated in flexible portals. Flexible execution frameworks on HPC (HTC, ensembles, VM images, etc.) Advanced scalable databases (KVP w/ mapreduce), ML at scale, in-situ analysis Big Data thumbnails, synopsis generation, Metadata automation, inferred provenance, De-noising, inter-dataset correlation, deep search, differential data sharing Automated agents for opportunistic data QA/QC, data citation, Social data, curation, community data management, attribution, Big Data reproducibility

RESTful Interface Circa 1955

190 REST APIs for Scientific Data and Computing

http://www.w3.org/community/hpcweb/

W3C Community and Business Groups
In situ processing*

These codes will be designed to run in situ on the same node as the simulation or on a tightly coupled co-processing node.

Data transfer: In memory or via ADIOS

Analysis
E.g. Merge tree
- Segment data and extract features of interest.
- Query particles and track features.

Visualization
E.g. Integrated volume, surface, and particle rendering
- Render raw data or segmented features.
- Render particles and trajectories.

Downstream tools
E.g. Statistical analysis of features, PCA
- Generate distributions, low-dimensional linear representations.

Post processing

These codes will be run as a post-process. They do not currently scale for in situ processing.

Analysis
E.g. MS-Complex
- Partition space using gradient flow.
- Cluster particles and trajectories.

Visualization
E.g. Integrated volume, surface, and particle rendering
- Render MS-complex.
- Render particles and trajectories clusters.

Downstream tools
E.g. Isomap
- Generate low-dimensional non-linear representations.

Simulation
E.g. S3D, GTC
Generate raw data.

Simulation
Time $t$

Time $t + \Delta t$

Feedback into Simulation:
E.g. regions for particle injection.


* Code for in situ processing can also be used in post processing.
Simula.on and Analysis Framework

Scientific Workflow envisioned

Beamline
User
Data Pipeline
HPC Storage and Compute
Science Gateway

New experiment

Prompt Analysis Pipeline

simulate
compare
measure
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.
New Models of Discovery

- Identify phenomena using machine learning
- New Science
- Validate models with experiments
- Fuse data with that of other scientists, disciplines
- Re-use and re-analyze previously-collected data
- Simulate with new models to understand data
- Discover relationships across data sets with sophisticated mathematical analyses
Identify Phenomenon using Machine Learning

• **TECA Toolkit today**
  - Automatic detection of cyclones, atmospheric rivers, and more
  - Analysis time years to minutes

• **Climate Analysis in 2031**
  - Machine learning for all events
  - Automatic metadata generation
  - Fusion of simulations, sensors, etc.
  - Real-time analysis and response
Connecting Data: Tools for Radical Scaling

- **Genomes to Life, KBASE**
  - Make genomics useful to biology
- **Microbes to Biomes**
  - Measuring and modeling the plant microbial biome
- **Quarks to Cosmos**
  - Frontiers bridge energy, intensity, and cosmos
- **Pixels to Knowledge**
  - Replace pixels with models, build kbases on models
  - Leverage repetition toward extreme structural resolution
- **Beamline to Browser**
  - Connect world’s Biggest Data instruments to the internet
- **Climate to Weather**
  - Couple world class global models to regional problem solving
- **Materials to machines**
  - Materials Project (replace materials design with search)
  - JCESR (batteries), JCAP (engineered sunlight-to-fuels tech)
  - Defects, functional electronics, nano-to-mesoscale

Big & Fast Filesystems
Powerful Flexible Computing
Advanced Analysis
Machine Learning
Ontologies, Ksystems
Databases/KVP
MapReduce
Software Defined Networking
High-throughput
Automated workflows
Filtering, De-Noise and Curating Data

Arno Penzias and Robert Wilson discover Cosmic Microwave Background in 1965

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
Re-Use and Re-Analyze Previously Collected Data

• Materials Genome Initiative
  – Materials Project: 4500 users 18 months!
  – Scientific American “World Changing Idea” of 2013 – what about 2031?

Unbounded computing requirements for simulation and analysis
Multi-modal analysis of Brain Connectivity

Analyze brain connectivity at multiple scales: From cells and regions to complex neural circuits.

• Improve understanding of brain pathology.
• Enable personalized treatment options.
Big Picture: Advancing Scientific Knowledge Discovery

- Knowledge management: collection, representation, storage, exchange, and sharing of large quantities of diverse information.
- Rapid information and knowledge-based response: decision-making mechanisms and support in near real-time.
- Data and knowledge fusion: integration of data and knowledge into consistent, accurate, and useful representation of the same or related real-world objects.
- Dynamic resourcing of data and information: discovery, allocation, and management mechanisms.
- Composition and execution of end-to-end scientific processes: configurable workflow methodologies spanning heterogeneous communities, applications, and environments.
- Human computer interaction: user interface, access, and interaction through computational environments and mechanisms.
- Trust and attribution: secure access, verification, and acknowledgement of contributions, curations.
How do we get there?
Extreme Data Scientific Facility (XDSF) Concept

Other data-producing sources

MS-DESI

ALS

LCLS

LHC

APS

JGI

ESnet

Energy Sciences Network
XDSF: Will bring scientists together with data researchers and software engineers
Software Defined Networking: A New Kind of Network (Especially for Science?)

- Internet is a black box—huge hidden complexity
- ESnet connects us to other labs and internet with novel “big science data” communications
- “Software Defined Network” demo’d in October:
  - Automatically adapts to large data transfers
  - Faster, cheaper, more flexible
  - Demo with Infinera. Brocade

See Greg Bell talk here at 4pm. Vern Paxson Plenary tomorrow
A view of the application space: Simulation and Data

<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>

NAP “Frontiers in Massive Data Analysis”
Data structures, Algorithms, Distributed Systems

- **Fastbit & Fastquery**
  - specialized compression and object-level search
  - bitmap indexing methods
  - Theoretically optimal and 10x-100x faster in practice

- **Tigres: Design templates for scientific workflows**
  - Explicitly support Sequence, Parallel, Split, Merge

L. Ramakrishnan, Valerie Hendrix, Daniel Gunter, Gilberto Pastorello, Ryan Rodriguez, Abdellilah Essari, Deb Agarwal
## Technology for Scientific Data

### Compute Intensive Arch

- **Goal:** Maximum Computational Density and local bandwidth for given power/cost constraint.

  Maximizes bandwidth density near compute

### Data Intensive Arch

- **Goal:** Maximum Data Capacity and global bandwidth for given power/cost constraint.

  Bring more storage capacity near compute (or conversely embed more compute into the storage).

  Requires software and programming environment support for such a paradigm shift
95% utilization, but the users wait
- Real-time analysis on streams
- Interactive access to data
Programming Challenge? Science Problems Fit Across the “Irregularity” Spectrum

Massive Independent Jobs for Analysis and Simulations
Nearest Neighbor Simulations
All-to-All Simulations
Random access, large data Analysis

... often they fit in multiple categories
The Programming Answer is Obvious...

More Regular

Message Passing Programming
Divide up domain in pieces
Compute one piece
Send/Receive data from others

MPI, and many libraries

More Irregular

Global Address Space Programming
Each start computing
Grab whatever / whenever

UPC, CAF, X10, Chapel, GlobalArrays
Computational Biologists buy large shared memory machines to assemble genomes.

For many problems (including metagenomics) these are not large enough.

Work by Evangelos Georganas, Jarrod Chapmanz, Khaled Ibrahim, Daniel Rokhsar, Leonid Oliker, and Katherine Yelick
End-to-end Computing for Science (Beam to bench)

Detector
- LCLS-II, ATLAS, Planck, K2, TEAM, PTF, FIB/SEM
- Synchrotrons & FELs
- Imaging Mass Spec
- Cryo EM
- Light and EM scopes

Network
- ESnet knows networks
- Built for science data
- Science DMZ, DTNs
- Making remote data local
- “Insight before a TB”
- SDN 1Tb/s in 2016?

Computer
- Petascale+ data analysis
- HTC & Realtime
- Science Gateways
- Community Databases
- NERSC8+, Exascale
Innovation from Data-driven Simulation Science

- BES sponsored NERSC HPC+HTC resources
  - 15M hours in 2012, 40M in 2013
  - 10X needed for nano-synthesis, MOFs
- Topically-focused allocations for national MGI program discovery challenges / target materials
- Advanced I/O for data analytics
- Data science focused HPC R&D

Reboot materials science as a collaborative HPC workflow, web based, durable data assets

*Ceder and Persson*
NERSC path to more data, science, and computing…

1. Meet the ever growing computing and data needs of our users by providing usable exascale computing and storage systems, transitioning SC codes to execute effectively on manycore architectures, and influencing the computer industry to ensure that future systems meet needs of SC.

2. Increase the productivity, usability and impact of SC’s data intensive science by providing comprehensive data systems and services to store, analyze, manage and share data.
Happy 40th NERSC!

Kathy Yelick
Associate Laboratory Director for Computing Sciences

David Skinner
NERSC Strategic Partnerships Lead

Lawrence Berkeley National Laboratory