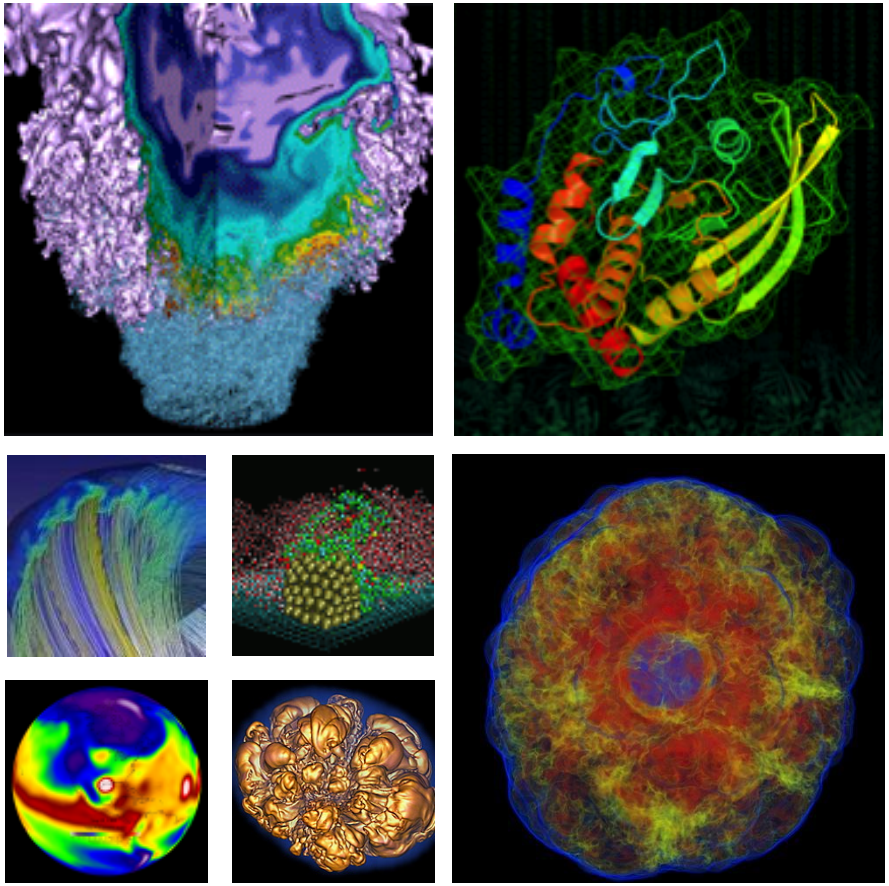
Towards a Light Source Data and Analysis Framework at NERSC



Jack Deslippe, Shane Canon, Abdelilah Essiari, Alexander Hexemer, Dula Parkinson, Simon Patton, **Craig Tull** + Many More

NERSC 40 YEARS at the FOREFRONT

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

The ALS Data Needs

Light source data volumes are growing many times faster than Moore's law.

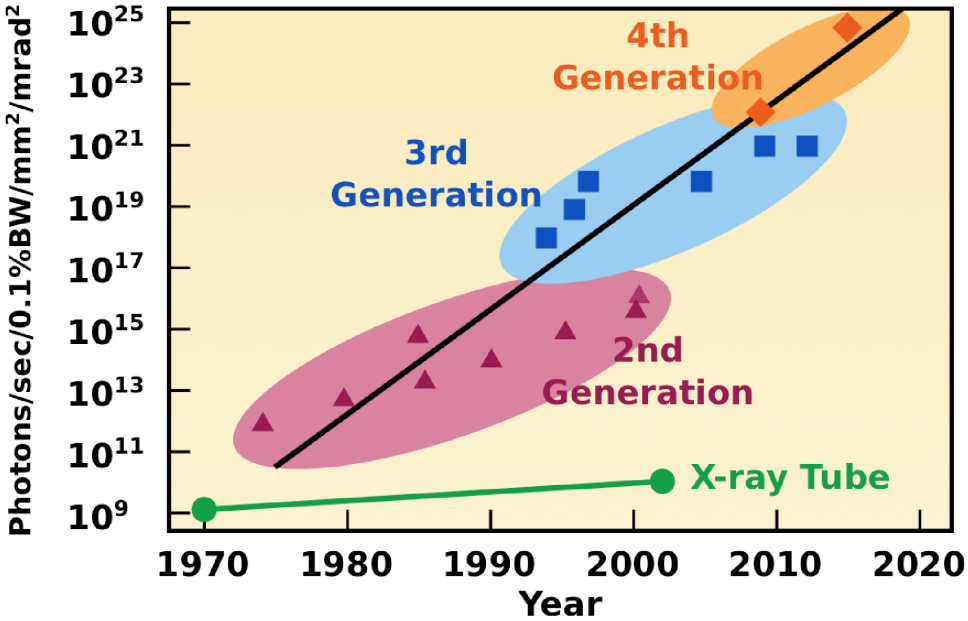
- Light source luminosity
- Detector resolution & rep-rates
- Sample automation

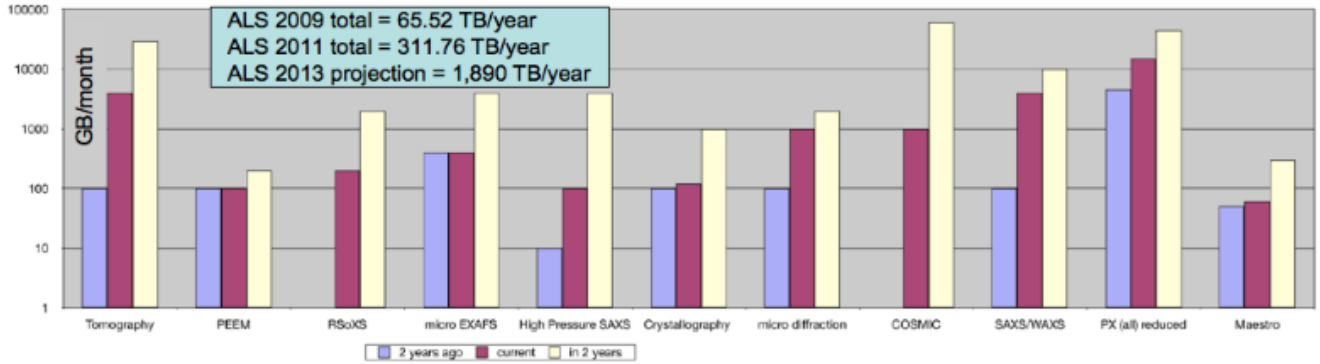
BES user facilities serve 10,000 scientists and engineers every year.

Mostly composed of many small groups with no SW experts.

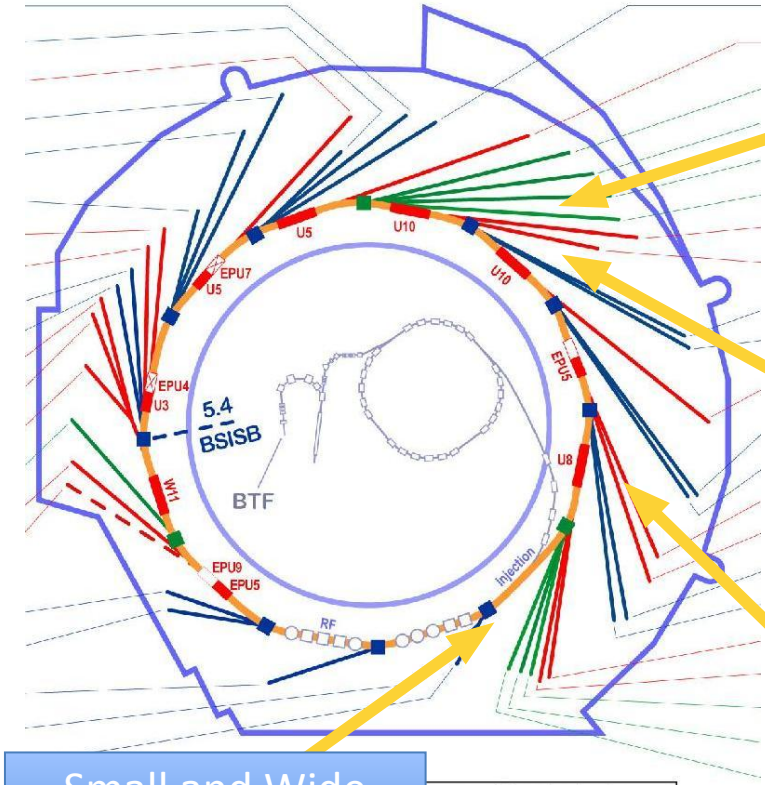
Simulation & analysis traditionally done on very small scale resources.

BES software landscape is largely ad-hoc and relies heavily on a limited number of experts to handle analysis.





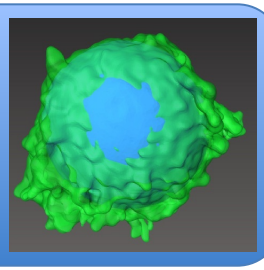
Diversity at the ALS



Comissioning

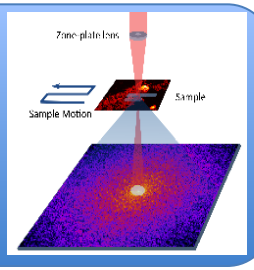
Tomography (Real Space)

Arec3d, QuantCT, CrunchFlow



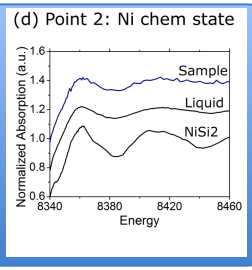
COSMIC Hybrid (Real and Reciprocal Space)

Ptychographic reconstruction



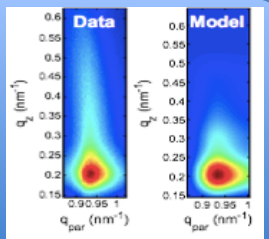
MicroXAS (Spectroscopy)

ShirleyXAS (MSD)
BerkeleyGW (NERSC)

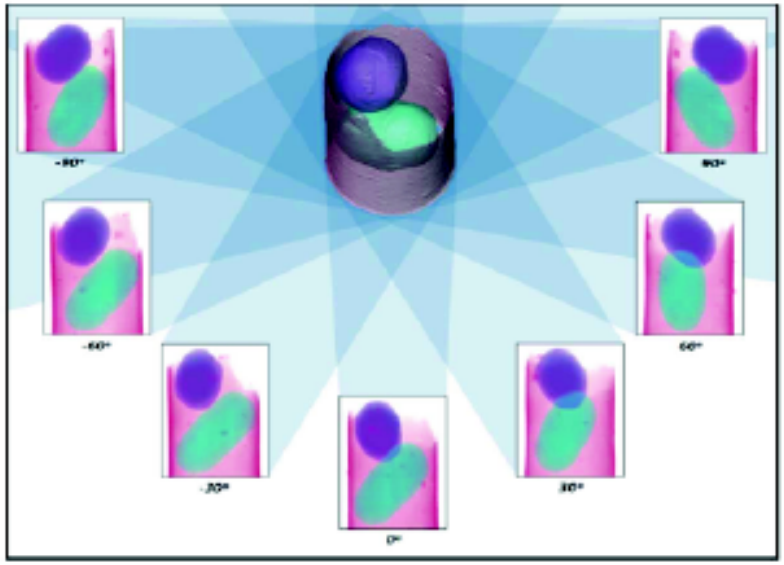


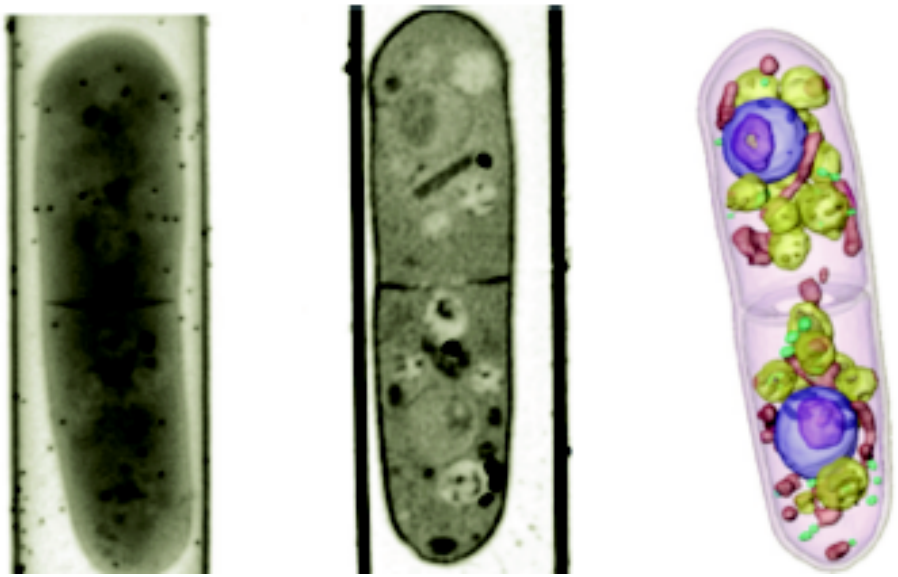
Small and Wide Angle Scattering (Reciprocal Space)

HipGISAXS/HipMC parallel Scattering



Tomography Beamline 8.3.2





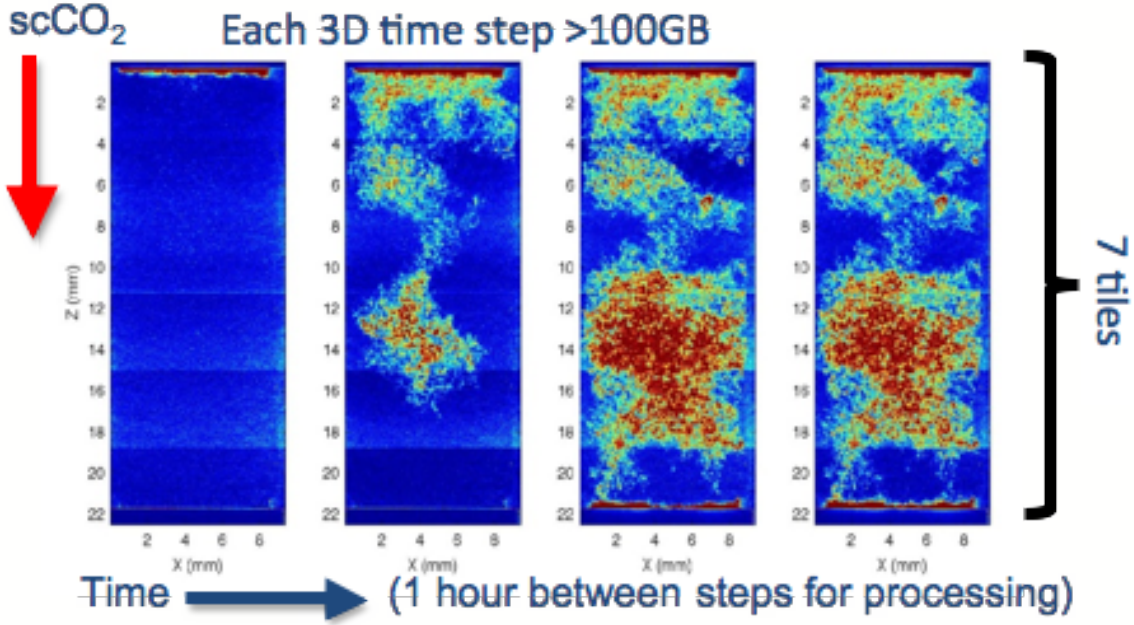
Raw Normalized Reconstructed

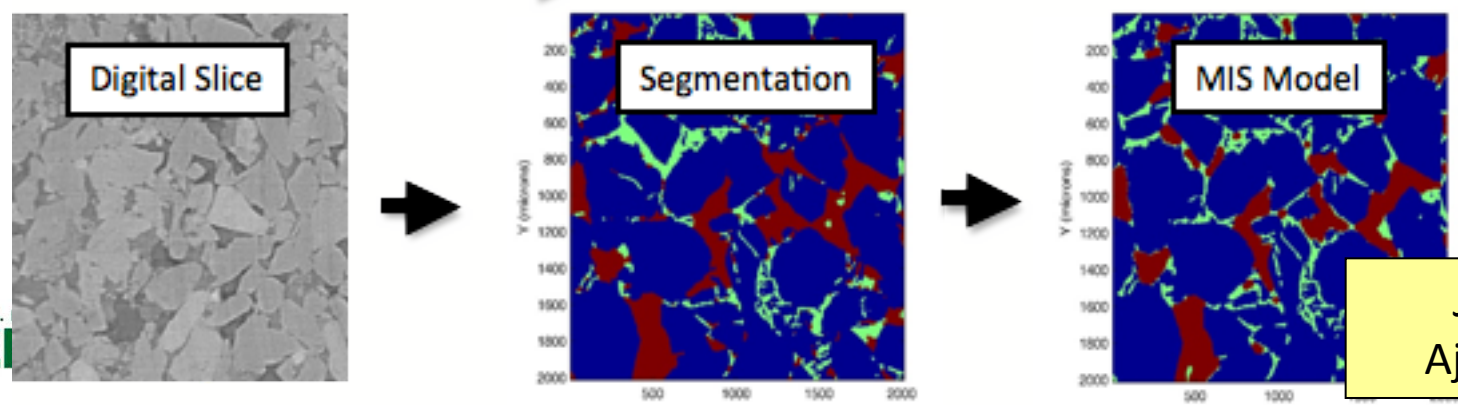
Measure X-Ray loss through a sample at various angles.
"Reconstruct" internal structure of sample from images.

Traditionally most users don’t achieve a reconstruction their data.

Geologic CO2 Sequestration

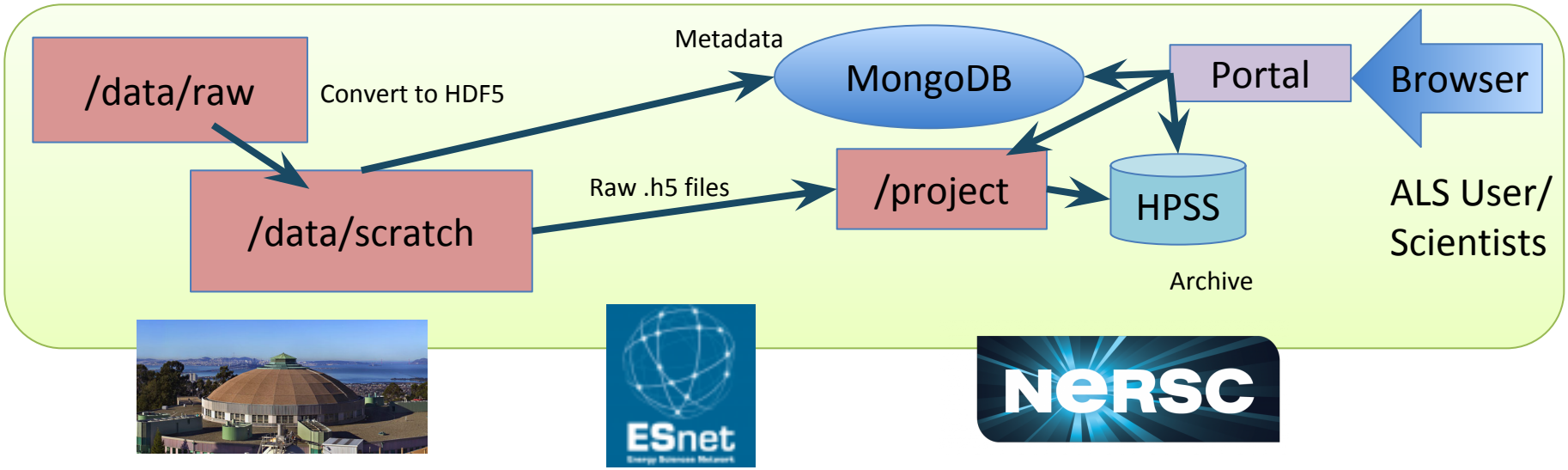
- 3-D beam-time feedback currently not available. Many users need this capability.
- Offline processing and simulation of large data sets impossible for many users.
- 3-D reconstruction often requires expert's direct intervention.



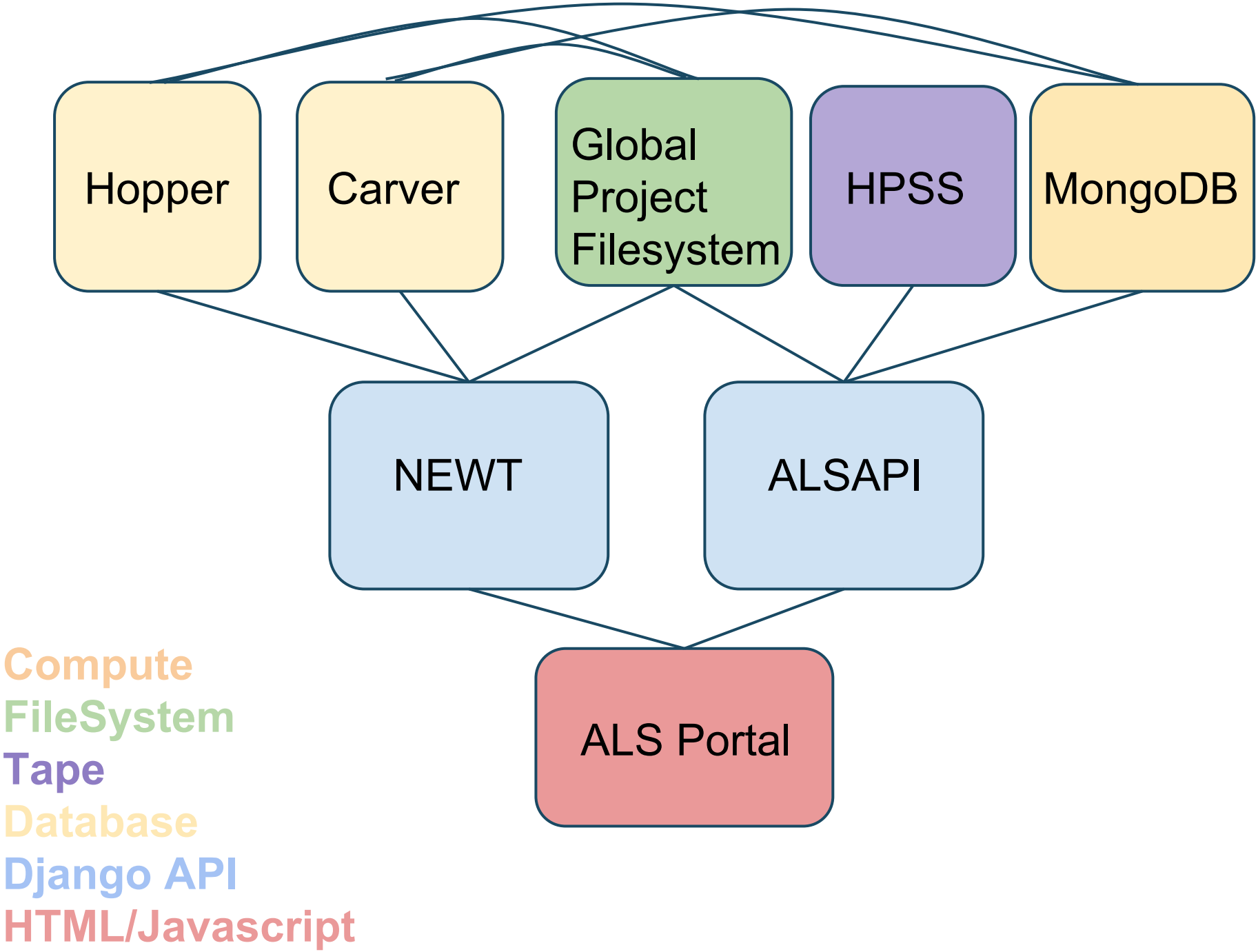


Jonathan Ajo-Franklin

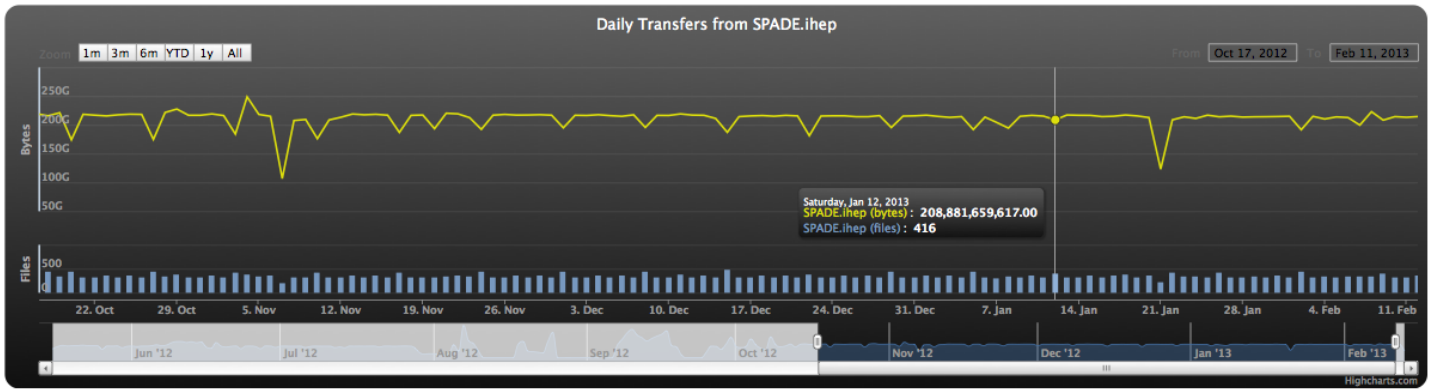
Data Pipeline Prototype in Action



- Prototype ALS Pipeline
  - Package Data into HDF format, existing data model standard (Nexus)
  - Capture critical metadata in HDF and Schema-less DB (MongoDB)
  - Automatically transfer to NERSC
  - Prototype web-access to allow ALS users to access data
- Investigating Methods to Store and Query Raw Sequence Data
  - Numbers of elements and query patterns exceed capabilities of existing off the shelf technologies.


Hopper
Carver
Global Project Filesystem
HPSS
MongoDB
NEWT
ALSAPI
ALS Portal
Compute
FileSystem
Tape
Database
Django API
HTML/Javascript

Data Transfer Pipeline



http://nest.lbl.gov/projects/spade/html/

- ALS Data Transfer Node (DTN): ESNet reference architecture
- Spade – Dayabay/IceCube production tool.
- Reliably transfer data from source to warehouse.
- Trigger actions at warehouse.
- Validate cksum and release source data file.
- GridFTP, Globus-Online, or scp transfer protocols
- Web and Mobile based monitoring and control

HDF5, MongoDB

HDF5 Advantages:

- Can compactly store multiple images in an organized way
- Can attach metadata to various datasets
- Highly performant IO for analysis

MongoDB Advantages:

- Flexible document store; different document fields for different beamlines
- Can create rich document-subdocument structure
- Can add/remove fields without worrying about schema
- Can index across multiple fields
- Outputs to JSON

The APIs

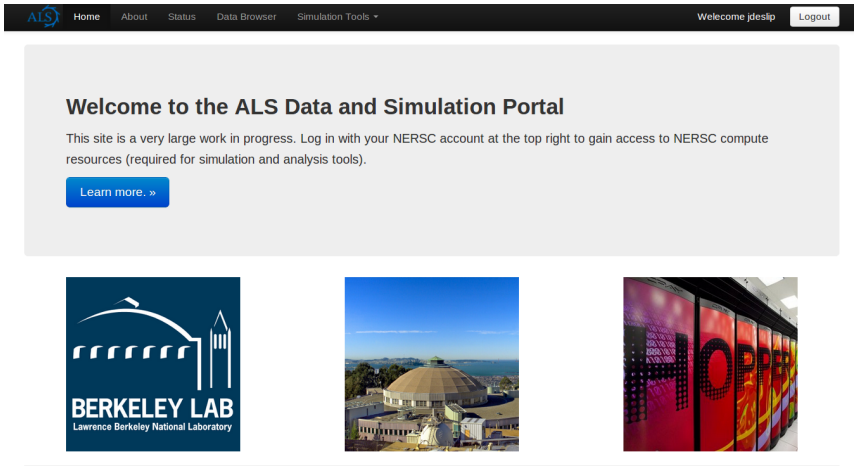
The "ALSAPI"

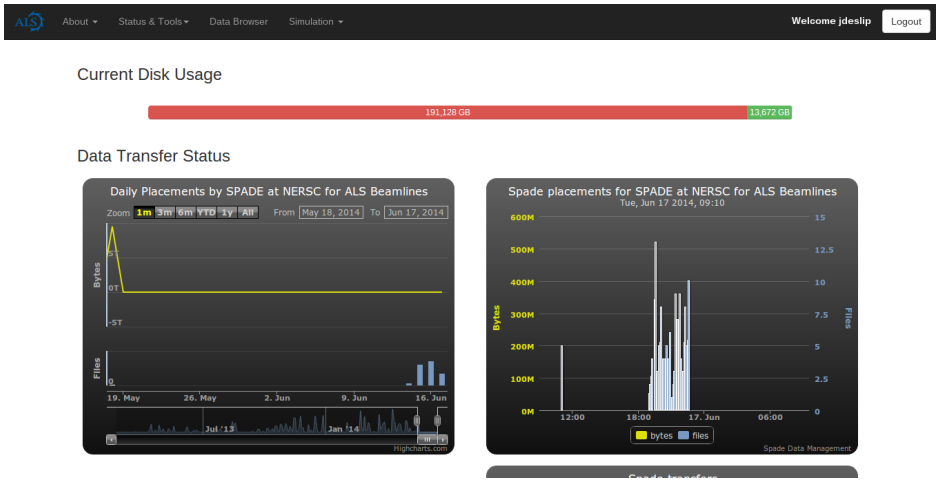
DJANGO App, runs on portal-auth.nersc.gov as "alsdata" user. Stages, converts, archives, retrieves and catalogs ALS data sets. Same "auth" as NEWT.

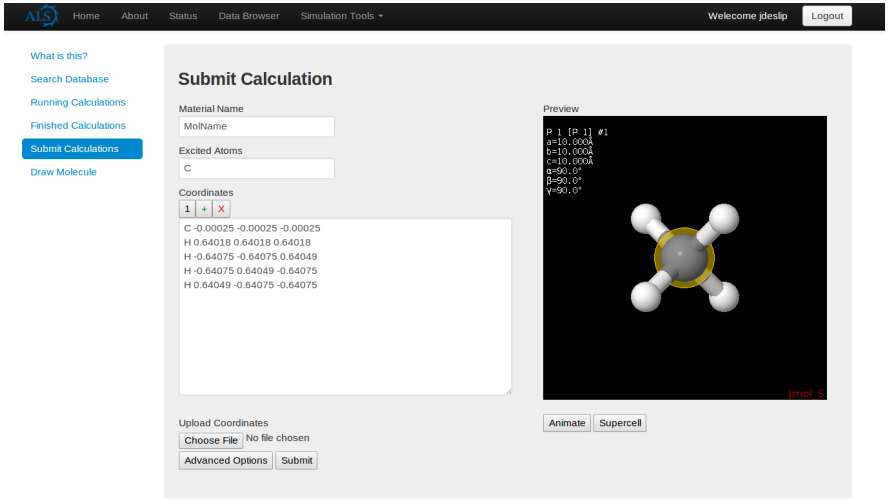
NEWT

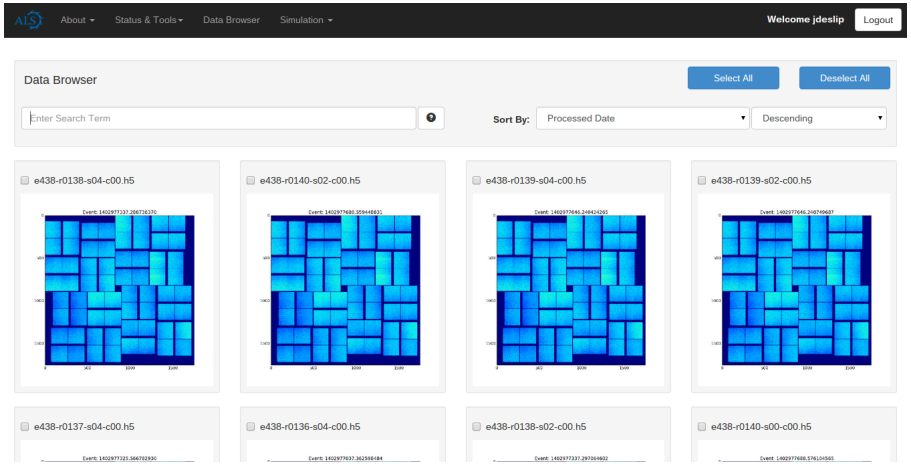
Users authenticate with NERSC credentials. Can run analysis and simulation software on Hopper and Carver. Results saved in user's area with option to share.

The Interface Features





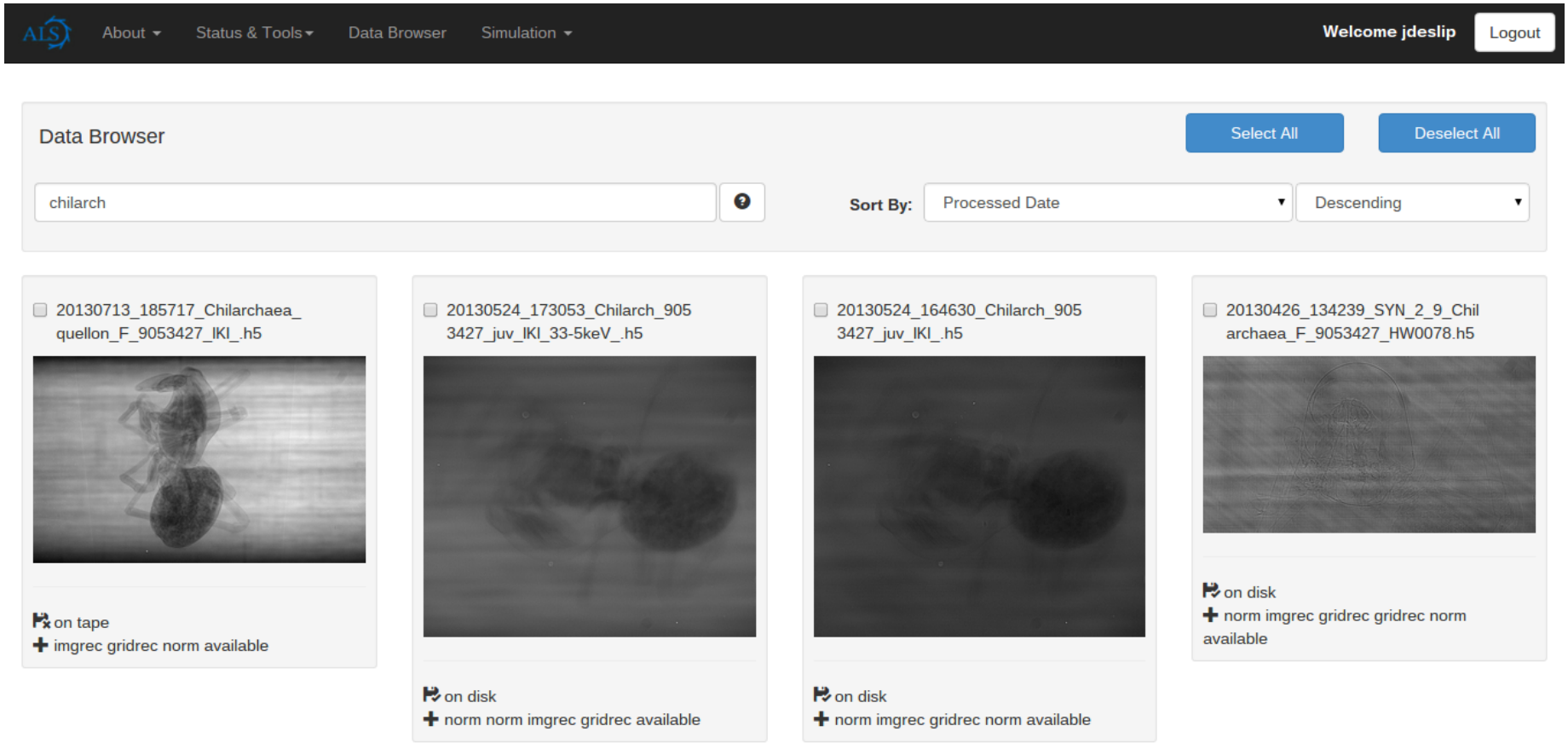




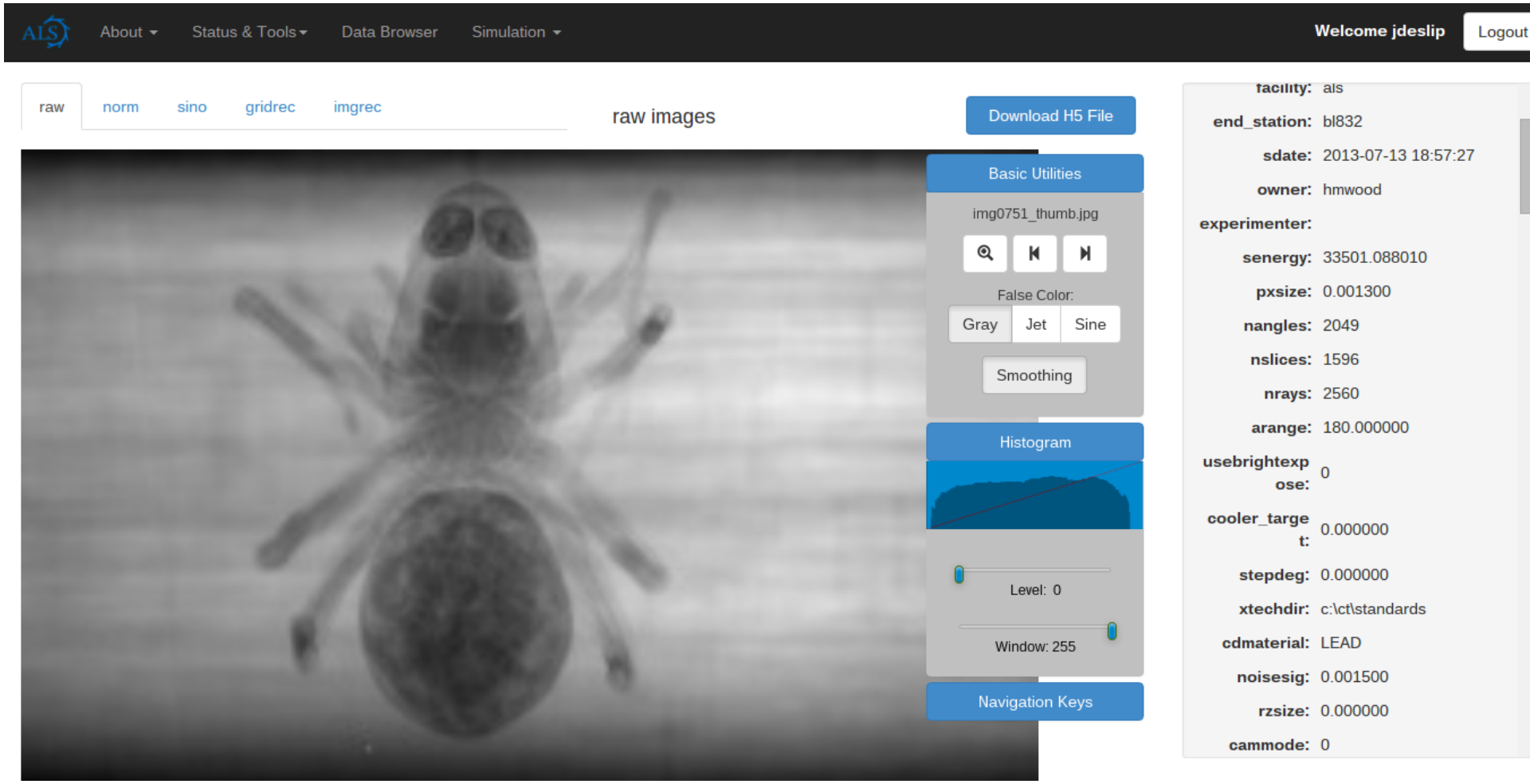
U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

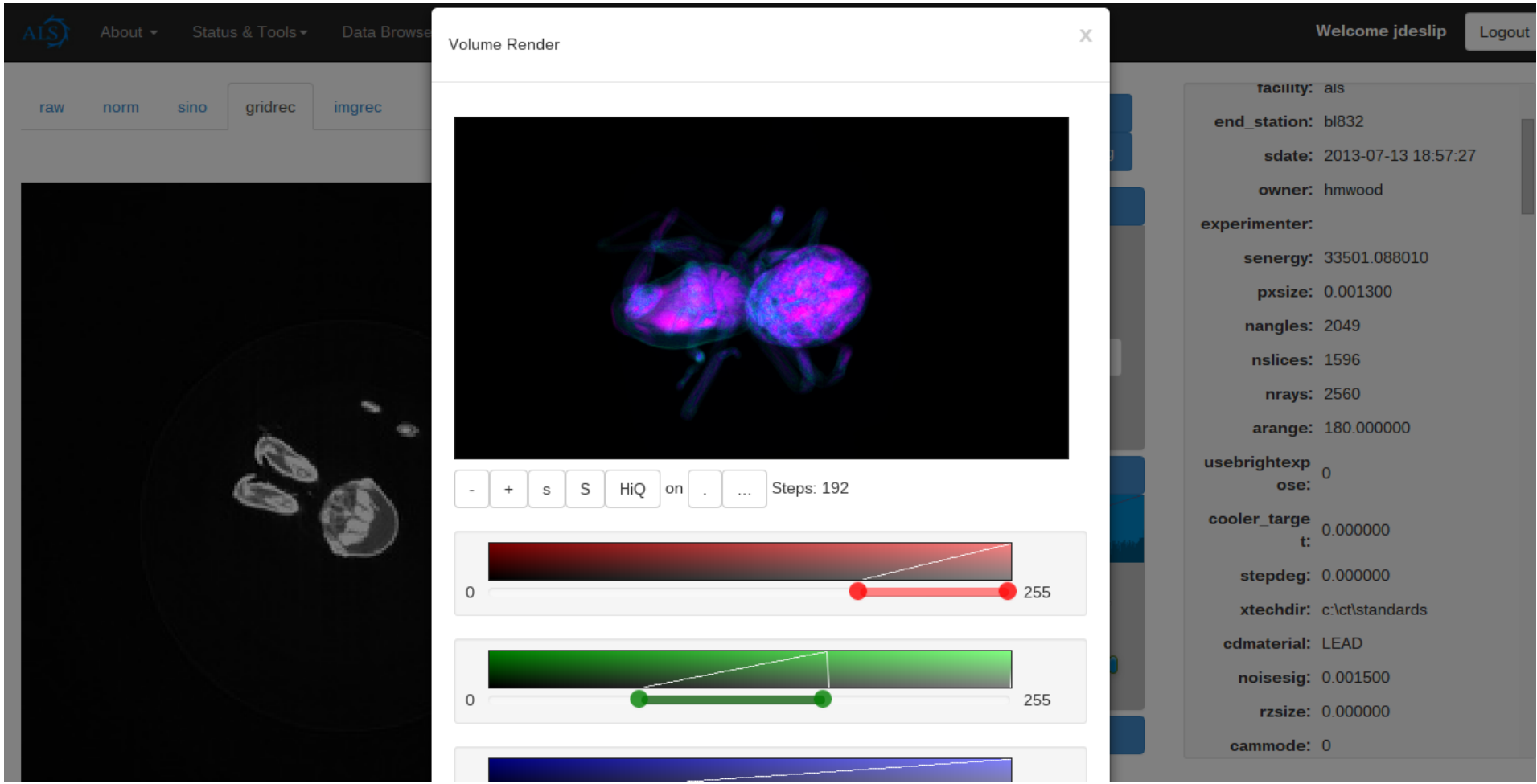
The Interface Features



The Interface Features



The Interface Features



Challenges and Future Directions

- Supporting multiple beamlines is a scaling challenge. Abstracting workflow, data transfer and providing a composeable analysis framework is necessary
- How guarantee real-time analysis at center with 99% usage and multi-day backlog. Killalble-Queue
- Turn-around on the order of minutes requires hiding communication costs - overlapping communication and computation.
- Standardization of data format across centers and beamlines. 3D vs 2D datasets within HDF5.
- Visualizing 100 GB datasets at full scale without jank.
- WAN distribution of data back

  to end-users

Total Datasets	89024
Total Data	477.9 TB
Total Raw Data	82.0 TB
Total Data Archived on Tape	477.9 TB
Total Data Staged on Disk	152.3 TB
Users with Data in System	114
Size of Mongo Database	39.25 GB

Contributors

- Craig E. Tull (CRD), A.Hexemer (ALS), B.Tierney (SND/ESNet), D.Prendergast (MSD), J.Deslippe (NERSC)

- Lab-wide LDRD (5 Divisions) funded FY 2013

  Computing Research Division, Advanced Light Source, Scientific Networking Division, Molecular Foundry, NERSC

- Richard Shane Canon [NERSC], Elaine Chan [ALS], Eli Dart [ESNet], Abedelilah Essiari [CRD], Dan Gunter [NERSC], Alexander Hexemer [ALS], Xiaoye Sherry Li [CRD], Filipe Maia [NERSC], Stefano Marchesini [ALS], Matthew Marcus [ALS], Inder Monga [ESNet], Dilworth Parkinson [ALS], Simon Patton [CRD], Eric Pouyoul [ESNet], David Prendergast [MSD], Lavanya Ramakrishnan [CRD], David Skinner [NERSC]

EXTRA SLIDES