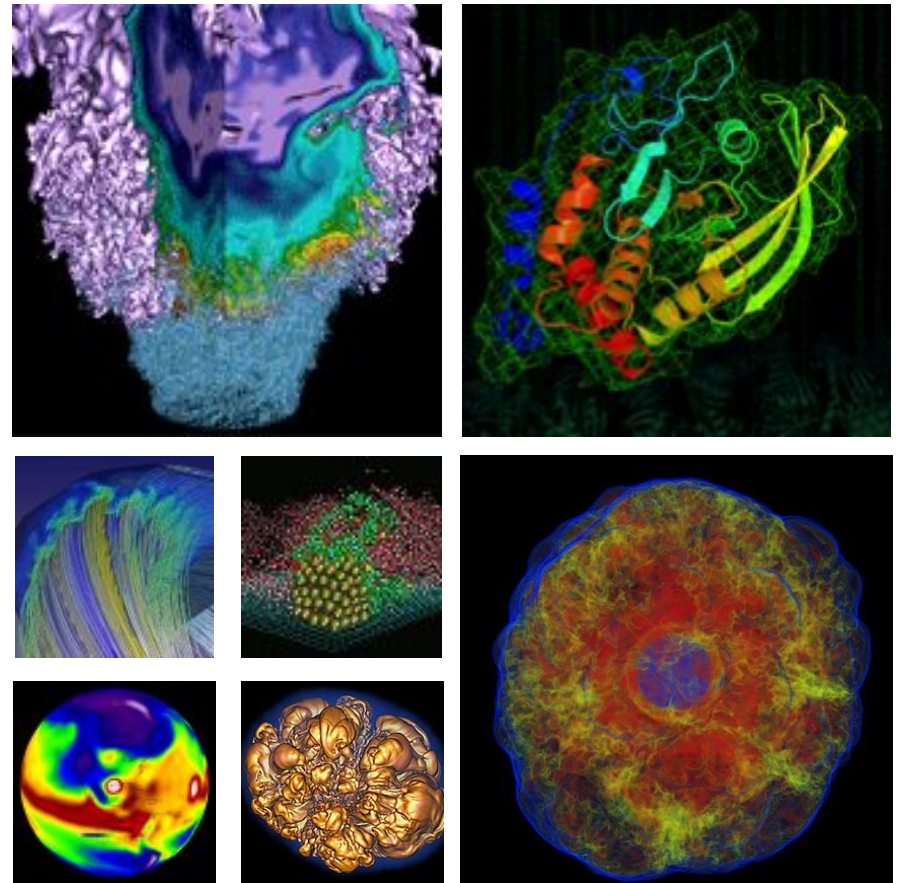


Science Gateways at NERSC



Annette Greiner, NERSC Data and Analytics Services

What is a Science Gateway?



- Web-based scientific resource for collaboration
- A form of computer-supported cooperative work
- Can be shared data, shared tools, may include social networking
- “value-added interfaces to access these shared resources” (sciencegateways.org)

- Traditional High Performance Computing assumes command-line expertise.

```
ssh carver.nersc.gov
qsub -I -q matgen_reg -l nodes=2:ppn=16 -l walltime=00:30:00
cd /project/projectdirs/matgen/vasp_test
mpirun -n 32 vasp
```

- The end-user scientist shouldn't have to be an old-school unix geek.
- Take advantage of rich visualizations and advanced user interfaces.
- Facilitate better sharing and collaboration.

Some Examples



Some Examples



Interfaces to HPC resources



Cr + O + Li

composition e.g. 1:1:1 Show unstable Show labels

Open Element: const μ

Mouse over dots for details

Compounds

Stable (11) Unstable (176)

Formula	Formation / Decomposition (Energy/Atom)	id
Cr	0	mp-90
Cr ₂ O ₃	-2.349	mp-19399
Cr ₅ O ₁₂	-1.847	mp-773920
CrO ₂	-2.054	mp-19177
Li	0	mp-135
Li ₂ CrO ₄	-2.107	mp-24920
Li ₂ O	-2.071	mp-1960
Li ₂ O ₂	-1.65	mp-841
Li ₃ CrO ₄	-2.191	mp-770632
LiCrO ₂	-2.399	mp-18793
O ₂	0	mp-12957

Some Examples



Data sharing across facilities/collaborations



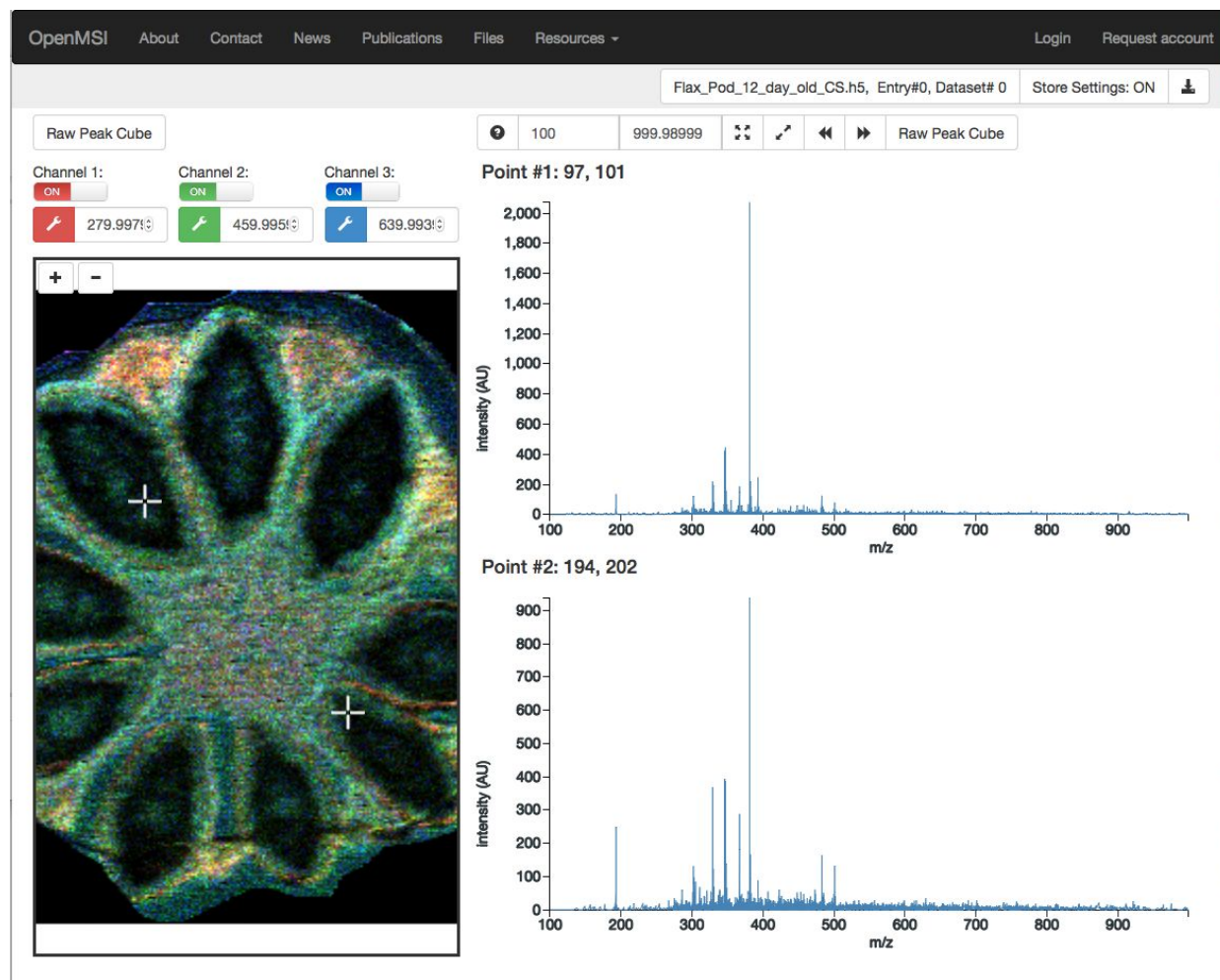
The screenshot shows the SPIN Data Browser interface. The top navigation bar includes 'SPIN', 'About', 'Status & Tools', 'Data Browser', and 'Simulation'. A user is logged in as 'Welcome agreiner' with a 'Logout' button. The main content area displays a grid of data items. Each item has a checkbox, a filename, a thumbnail image, and a list of available processing options.

File Name	Thumbnail Description	Processing Options
<input type="checkbox"/> 20150307_185312_HypScan_595_x00y01.h5	Micrograph of a textured surface with a central vertical feature.	<input type="checkbox"/> on tape + norm sino gridrec imgrec rc-gridrec rc-imgrec available
<input type="checkbox"/> 20150307_185312_HypScan_595_x00y00.h5	Micrograph of a textured surface with a central vertical feature.	<input type="checkbox"/> on tape + norm sino gridrec imgrec rc-gridrec rc-imgrec available
<input type="checkbox"/> 20150214_140815_beamstop_tes_t_150214140813725.h5	[Redacted]	[Redacted]
<input type="checkbox"/> 20150214_214834_beamstop_tes_t_150214214833626.h5	[Redacted]	[Redacted]

Some Examples



Interactive tools



Some Examples



Rich visualizations and UIs



KBase fruitfly
Created by: Annette Greiner (agreiner)

kernel share save

Analyze Narratives Jobs

DATA

- DM_tree v1**
Tree
5 minutes ago
- my_features v1**
FeatureSet
62 minutes ago
- Drosophila_melanogaster__Fruit...**
Genome: Drosophila melanogaster (Fruit fly)
66 minutes ago

APPS & METHODS

- Domains** Annotate Domains in a Genome v1.0.0
- RAST** Annotate Microbial Contigs v1.0.0
- RAST** Annotate Microbial Genome v1.0.0
- Plant SEED** Annotate Plant Coding Sequences with Metabolic Functions v1.0.0

Status Console Log

Job Id: 57981820e4b07760c9df64d0

Status: Completed

Insert Genome Into Species Tree (method output) 19:13:43, 7/26/2016

Phylogenetic tree visualization showing relationships between various species, including Thermococcus, Sulfolobus, Methanococcus, Methanobrevibacter, Candidatus Methanosphaera, Methanoculleus, Methanosarcina, Natronomonas, Halogeometricum, Haloferax, and Drosophila melanogaster. Bootstrap values are shown at the nodes.

Three Levels of Collaboration



Self-Service

Initial Consulting

Immersive Engagement

User-Built Gateways



qcd.nerisc.gov

legacysurvey.org

cxidb.org



Building Your Own Gateway



What level of service do you need? (We are 8x5)

Do you need to work with a programmer? (Domain Expertise + CS expertise)

What kind of data do you have? (HDF5, Database, Files)

Do you have long running computations? (> web time)

Would you benefit from an API? (Making data available in a structured format)

Building Your Own Gateway



The quick way:

World-readable files in /project/projectdirs/myproject/www

/project/projectdirs/myproject/
www/index.html

<http://portal.nersc.gov/project/myproject/index.html>

```
<!DOCTYPE html>
<html lang="en">
  <head>
    <meta charset="utf-8">
    <title>title</title>
  </head>
  <body>
    <h1>Hello, world!</h1>
  </body>
</html>
```



Django and Flask (Python) are popular web frameworks

Python has an extensive set of scientific libraries like NumPy, SciPy, H5PY

See also: PHP, web2py, etc.

PyDAP for serving HDF5 data

Jupyter Notebook, R-Studio for Interactive Computing

Globus, NEWT



We've found that scientists often want programmatic access to data

e.g. Materials Project: Give me property X for all materials with Li and O so that I can pass it through my own codes

Lesson – make your data available through an API and people will start to do new and innovative things

Web API Example



GET <https://www.materialsproject.org/rest/v1/materials/Fe2O3/vasp/energy>

```
{
  "created_at": "2013-03-17T09:14:58.158081",
  "valid_response": true,
  "version": {
    "pymatgen": "2.5.4",
    "db": "2013.02.25",
    "rest": "1.0"
  },
  "response": [{
    "energy": -132.33005625,
    "material_id": 542309
  }, {
    "energy": -66.62512425,
    "material_id": 24972
  }],
  "copyright": "Copyright 2012, The Materials Project"
}
```

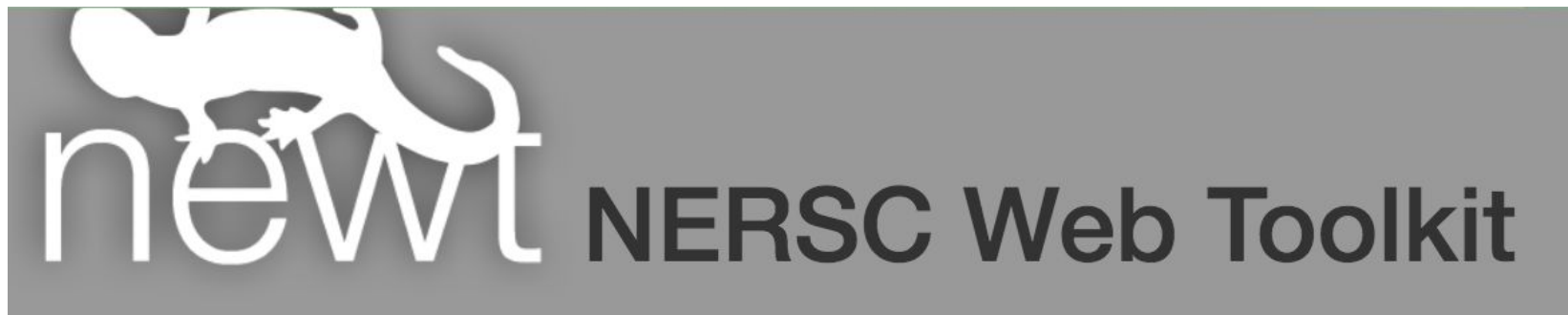
REST API for HPC resources

NERSC over HTTP

Jobs, Files, Commands, Status, Accounting, etc.

Makes it very easy to build web applications that
can interface with NERSC

<https://newt.nersc.gov>



NEWT Example



GET <https://newt.nersc.gov/newt/status/>

```
[{
  "status": "up",
  "system": "cori"
}, {
  "status": "up",
  "system": "edison"
}, {
  "status": "up",
  "system": "pdsf"
}, {
  "status": "up",
  "system": "genepool"
}, {
  "status": "up",
  "system": "archive"
}]
```

Dogfooding



My **NERSC**

- agreiner
- Dashboard
- Queues
- Center Status
- File Browser
- Jobscript Generator
- Completed Jobs
- My Tickets
- Data Dashboard
- NX Desktop
- Changelog
- NERSC Homepage

Data Dashboard

Showing disk space and inode usage for project directories at NERSC to which you have access as PI, PI proxy, or user (includes /project, /projecta, and /projectb/sandbox)

als directory in **/project** as of Thu Feb 16 2017 03:26:33 GMT-0800 (PST) [Toggle Details](#)

Usage as Percent of Allocations

66% of space

78% of inodes

CAL directory in **/project** as of Thu Feb 16 2017 03:26:33 GMT-0800 (PST) [Toggle Details](#)

Usage as Percent of Allocations

81% of space

3% of inodes

coe directory in **/project** as of Thu Feb 16 2017 03:26:33 GMT-0800 (PST) [Toggle Details](#)

Usage as Percent of Allocations

4% of space

0% of inodes

cosmo directory in **/project** as of Thu Feb 16 2017 03:26:33 GMT-0800 (PST) [Toggle Details](#)

What Next?



Read the Docs:

<http://www.nersc.gov/users/science-gateways/>

What Next?



Contact the science gateways team



Shreyas Cholia
Access, Management



Rollin Thomas
Astrophysics, Cosmology
Analytics, Access



Annette Greiner
Biomedical Sciences
Access, Visualization

consult@nersc.gov

The logo for the National Energy Research Scientific Computing Center (NERSC) is displayed in a large, white, sans-serif font. The letters are set against a dark blue background with a radial pattern of lighter blue lines emanating from behind the text, creating a sunburst or starburst effect. The entire graphic is contained within a dark blue rounded rectangular border.

NERSC

**National Energy Research
Scientific Computing Center**