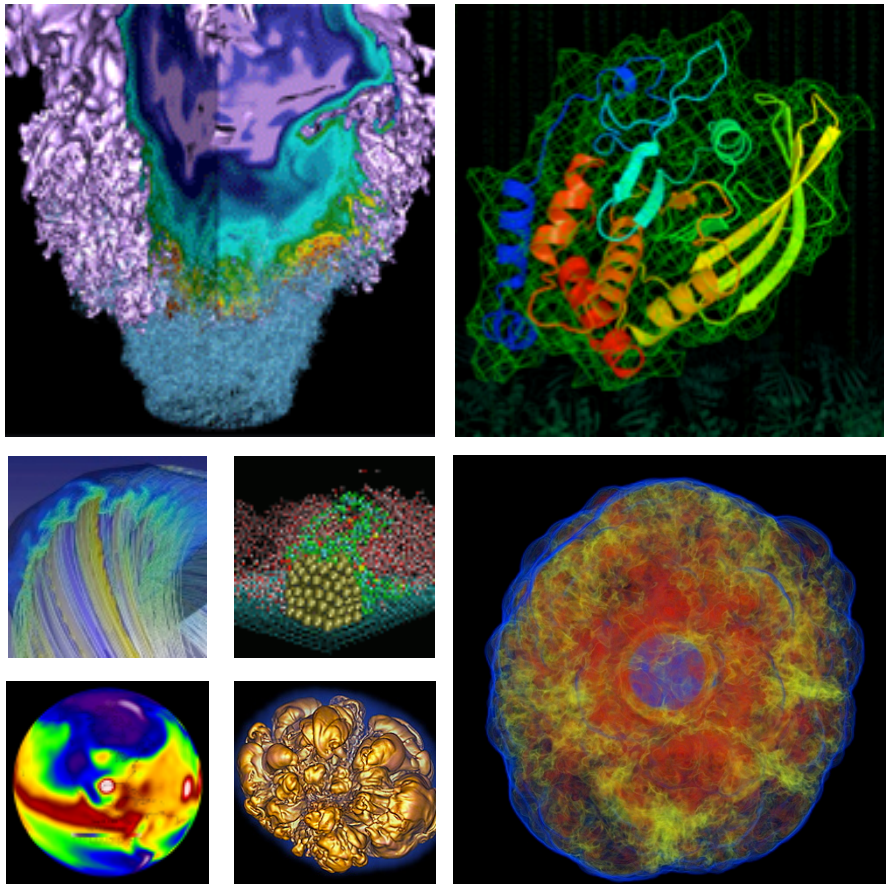
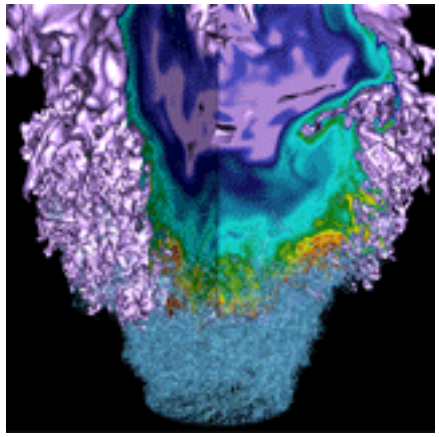
Moving And Sharing Data at NERSC

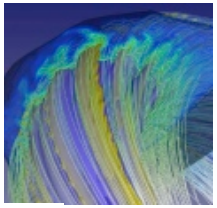


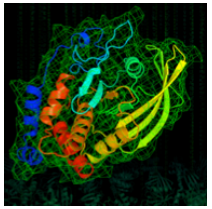
Shreyas Cholia
scholia@lbl.gov
Data and Analytics Services

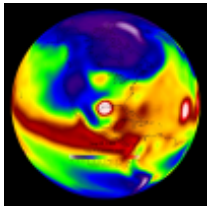
New User Training 2016
March 21th, 2016

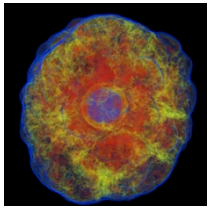
Data Transfer Nodes

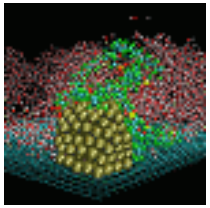












NERSC 40 YEARS at the FOREFRONT

Dedicated Data Transfer Systems: Data Transfer Nodes

- **Data Transfer Nodes (DTN) are servers dedicated to data transfer at NERSC.**
  - Currently 4 Nodes → dtn[01-04].nersc.gov

- **DTN features**
  - High bandwidth network interfaces
  - Access to global NERSC file systems
  - Tuned for efficient data transfers
  - Tuned for transferring large volumes of data between NERSC and other major facilities (ORNL, ANL etc.)
  - Can also move data between NERSC file systems and HPSS

➢ **Use the DTNs if you want to move large volumes of data in and out of NERSC**

DTN Login Access

- **All NERSC users have login access**
  - NERSC Users (non-JGI):
    - `ssh dtn01.nersc.gov` (or `dtn02`)
  - JGI Users:
    - `ssh dtn03.nersc.gov` (or `dtn04`)

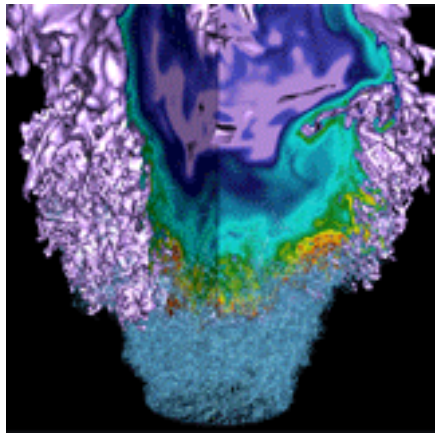
- **Familiar module environment**
  - `module avail`
  - Limited software deployment for data transfer needs

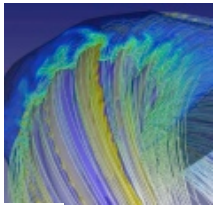
DTN File Systems

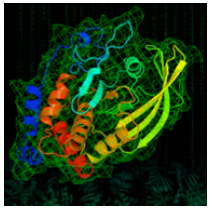
- **Global File Systems are available**
  - `/global/homes`
  - `/global/project`
  - `/global/projecta`
  - `/global/projectb`
  - `/global/dna`
  - `/global/seqfs`
  - `/global/common`

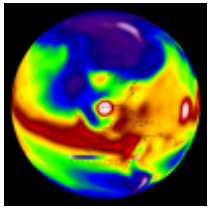
- **Currently excludes system-specific file systems**
  - e.g. /scratchX on Cori & Edison

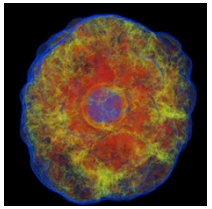
Moving Data

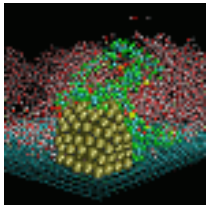










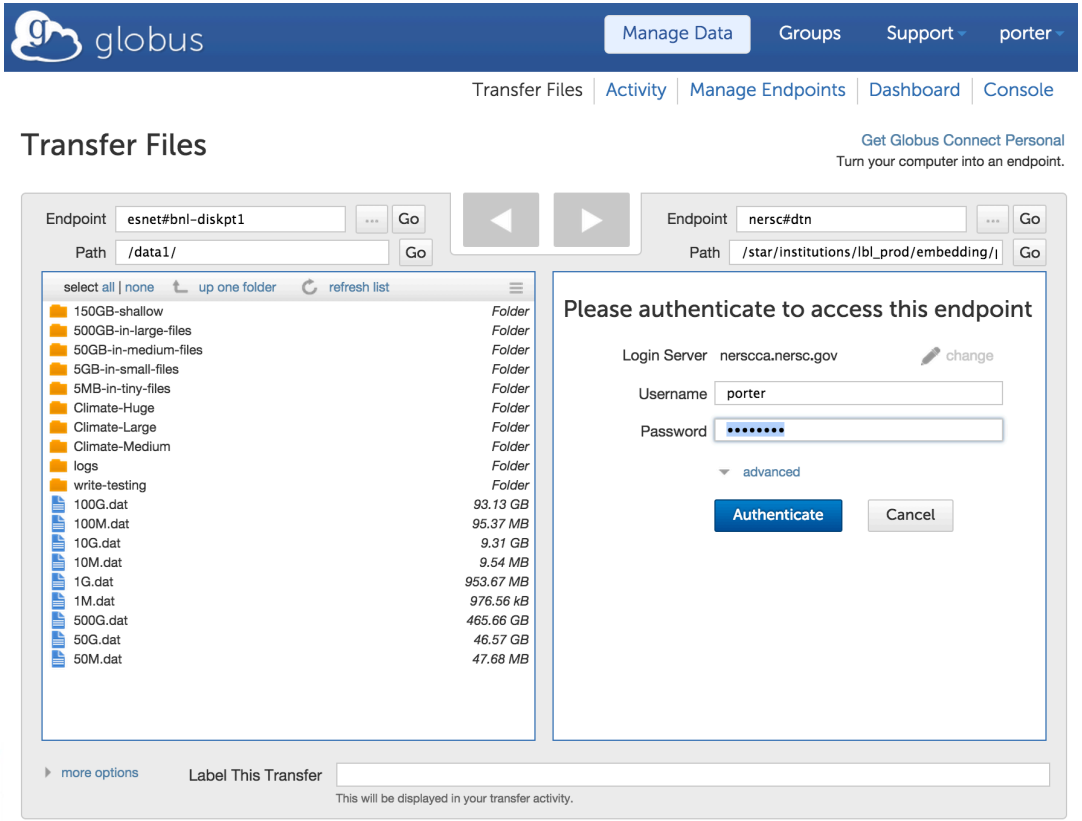


Transfer Tools

- **Globus: http://globus.org**
  - Reliable transfers of large data sets between sites or systems
- **scp**
  - copying individual files and directories
- **GridFTP (globus-url-copy)**
  - high performance CLI w/ grid certificates
- **bbcp**
  - high performance CLI tool w/ minimal setup
- **hsi/htar**
  - for data transfer into & out of HPSS

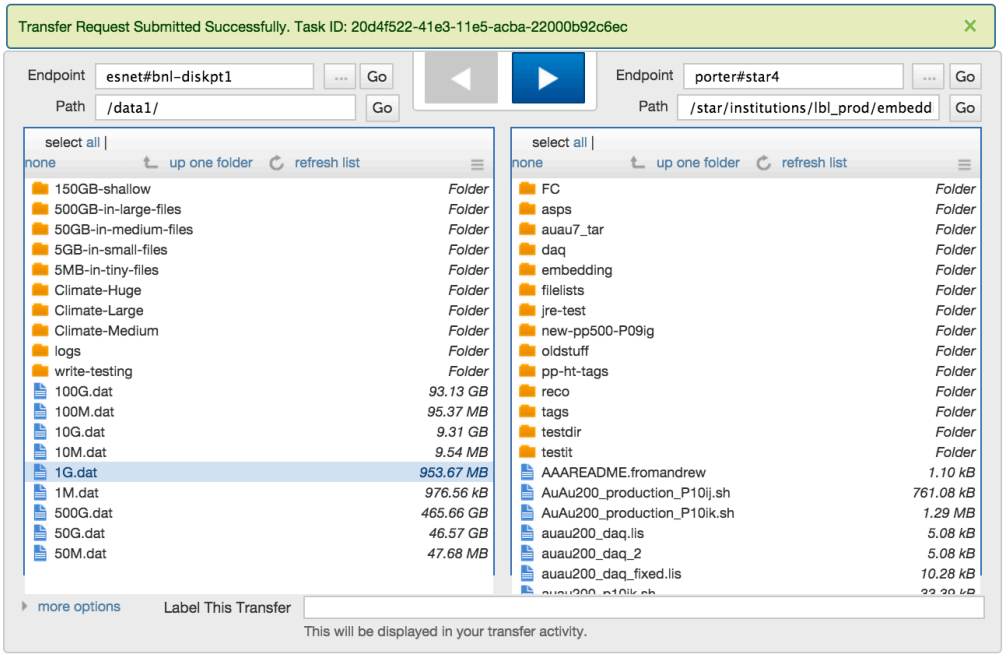
Globus.org

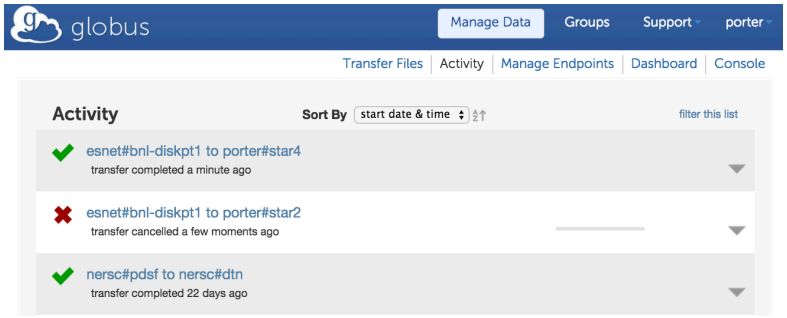
- **Managed 3rd party transfers** https://www.globus.org/
  - Web based interface, currently requires Globus Account
- **Type `nersc#dtn` endpoint in the transfer window**
  - Activate with NERSC username & password



Globus.org Continued

- **Use point & click interface to submit**
  - Transfer happens in the background & retries on failure
  - You can status on Activity panel
  - Email notification on completion





Globus Notification
To: Jeff Porter
SUCCEEDED - 20d4f522-41e3-11e5-acba-22000b92c6ec

TASK DETAILS
Task ID : 20d4f522-41e3-11e5-acba-22000b92c6ec
Task Type : TRANSFER
Status : SUCCEEDED
Is Paused : No
Request Time : 2015-08-13 17:45:47Z
Deadline : 2015-08-14 17:45:46Z
Completion Time : 2015-08-13 17:46:08Z
Total Tasks : 1
Tasks Successful : 1
Tasks Expired : 0
Tasks Canceled : 0
Tasks Failed : 0
Tasks Pending : 0
Tasks Retrying : 0
Command : API 0.10 go
Label : n/a
Source Endpoint : esnet#bnl-diskpt1
Destination Endpoint : porter#star4
Sync Level : n/a
Data Encryption : No
Checksum Verification : Yes
Delete : No
Files : 1
Files Skipped : 0
Directories : 0
Expansions : 0
Bytes Transferred : 1000000000
Bytes Checksummed (Sync): 0
Effective MBits/sec : 380.952
Faults : 0

- **Globus Connect allows a private endpoint on your laptop**
  - https://www.globus.org/globus-connect

Other tools

- **Command line tools for use directly from the DTN nodes or from a remote node**
  - scp
  - GridFTP (globus-url-copy)
    - http://tinyurl.com/nersc-gridftp
  - bbcp
    - http://tinyurl.com/nerscbbcp

Tool	Throughput
scp	140 Mbps (17.5 MB/s)
HPN patched scp, 1 disk	760 Mbps (95 MB/s)
HPN patched scp, RAID disk	1.2 Gbps (150 MB/s)
GridFTP, 1 stream, 1 disk	760 Mbps (95 MB/s)
GridFTP, 1 stream, RAID disk	1.4 Gbps (175 MB/s)
GridFTP, 4 streams, RAID disk	5.4 Gbps (675 MB/s)
GridFTP, 8 streams, RAID disk	6.6 Gbps (825 MB/s)

- **Detailed instructions, syntax etc.**
  - http://tinyurl.com/nerscdtn

http://fasterdata.es.net/data-transfer-tools/

scp - Secure Copy

- **Uses SSH under the covers**
- **Good for "small" (~100s of MB)**
- **Simple cp-like syntax**
  - `scp localfile user@host:remotefile`

- **At NERSC, we deploy high performance modifications (HPN SSH) to get better throughput**

HPSS

- **Backup or archive your data with HSI and HTAR**
  http://www.nersc.gov/users/data-and-file-systems/hpss/getting-started/
  (http://tinyurl.com/nerschpss)

- **Login to DTN node and use hsi/htar**
  - HSI for individual files and conditional access
  - HTAR for aggregation & optimization of storage/archival

- **Can also use Globus Online:**
  - External endpoint ←→ nersc#hpss
  - nersc#dtn ←→ nersc#hpss
  - But individual files only, does not support htar

General Tips

- **Use DTNs as dedicated data transfer servers**
  - Tuned for WAN transfers
  - Fast network (ESnet), optimized configuration
    - See http://fasterdata.es.net/ for more discussion
  - Dedicated support for data transfer
- **Don't use DTN nodes for non-data transfer purposes**
- **Use Globus for large, automated or monitored transfers**
- **scp is fine for smaller, one-time transfers (<100MB)**
  - Globus is also fine for small transfers
- **Plain "cp" or "rsync" is still used for local transfers**

Performance Considerations

- **Performance is often limited by the remote endpoint**
  - Not tuned for WAN transfers
  - Have limited network (~1Gb/sec) link.
  - These will lower performance < 100 MB/sec.

- **File system contention may be an issue**
  - For example, don't use your $HOME directory!
  - Instead use /project, $SCRATCH or $GSCRATCH

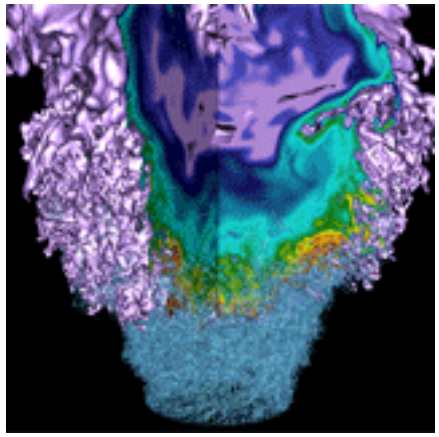
- **If you don’t think you are getting the performance you expect, let us know: consult@nersc.gov**

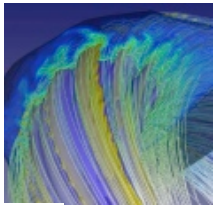
For more information

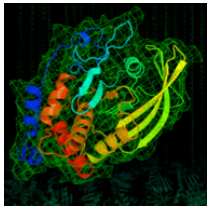
- **General DTN info**
  - http://www.nersc.gov/systems/data-transfer-nodes/
- **Data transfer info**
  - http://www.nersc.gov/users/data-and-file-systems/transferring-data/

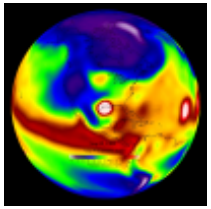
    (or http://tinyurl.com/nerscdtn)
- **Feedback / Problems?**
  - consult@nersc.gov
- **Globus Support**
  - https://www.globus.org/support/

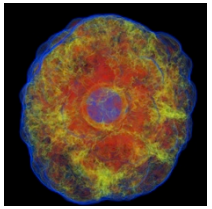
Uploads from Outside Users

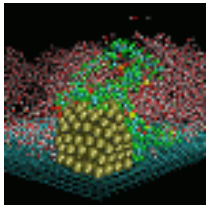












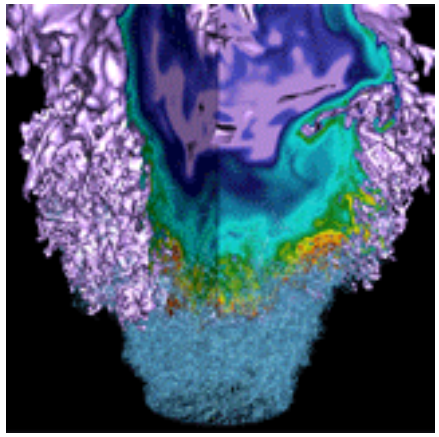
Anonymous FTP Upload Service

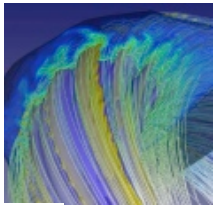
- **For external collaborators to be able to send data to NERSC users**
- **Create a temporary FTP account to upload files that will be delivered to a NERSC user**
- **NERSC user can retrieve the file:**

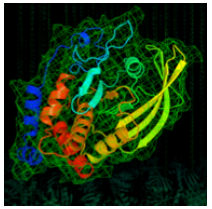
```
take -u ftpup <file_name>
```

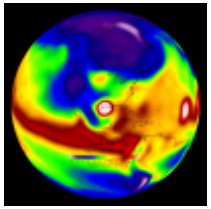
- **https://www.nersc.gov/users/storage-and-file-systems/transferring-data/nersc-ftp-upload-service**

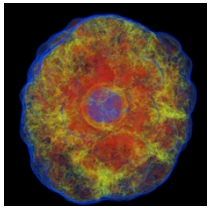
Sharing Data Within NERSC

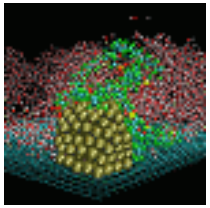












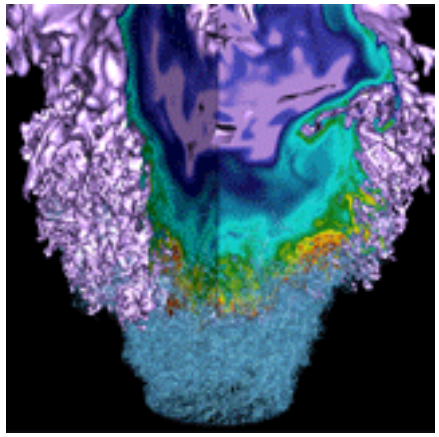
The /project directory

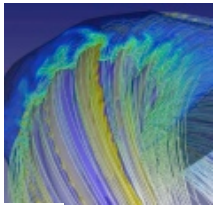
- Your project will have a shared project directory at /global/project/projectdirs/<dirname>
- Shared group permissions for your project repo
- https://www.nersc.gov/users/storage-and-file-systems/sharing-data/

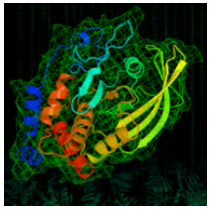
Give/take

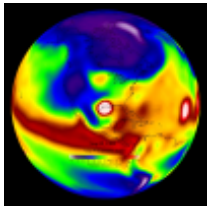
- **Give a file to a user:**
  - give -u <recipient-username> <file-or-directory>
  - This sends is into a holding area for pickup and notifies the recipient over email
- **Accept a file from a user**
  - take -u <sender-username> (-d <destination folder>) <filename>
  - Picks up the file (use –a for all files)

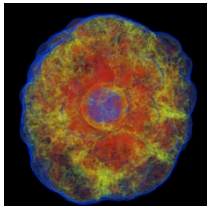
Science Gateways: Sharing Data over the Web

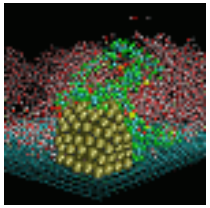








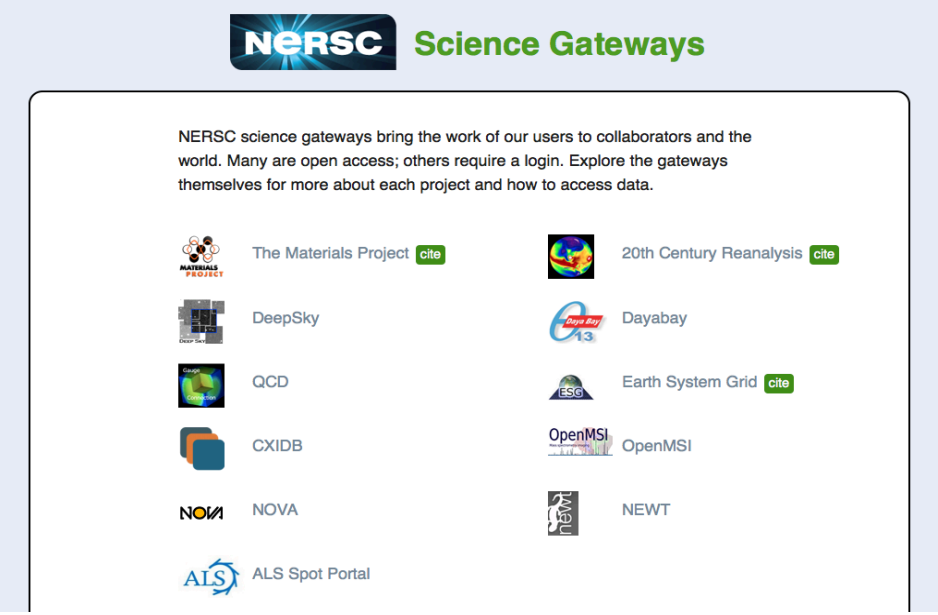




NERSC 40 YEARS at the FOREFRONT

NERSC Science Gateways

- **Web portals that allow you to interface with your data and computation at NERSC**
- **Interfaces built around your science**
- **Science-As-A-Service**



Services

- **Simple data publishing capabilities**
- **Rich web interfaces and complex portals**
- **Backend databases and message queues**
- **Interactive Analysis Tools over the Web**
  - iPython Notebooks and R Studio
- **NEWT API to access NERSC resources**

Publish Data On the Web

- **Every repo now has a project directory**
- **Trivial to wire up your project directory to make it accessible over the web**
- **Create a file in your www directory**
  - mkdir /global/project/projectdirs/<yourproj>/www
  - cd /global/project/projectdirs/<yourproj>/www
  - vi index.html

  <html>Hello World</html>
- **Make sure all the above files and directories are world readable**
  - chmod 775 /global/project/projectdir/<yourproj>/ etc.
- **Voila:**
  - http://portal.nersc.gov/project/<yourproj>/

Build Full Stack Web Applications

- **Build full stack web applications for your science at NERSC**
  - **Python/Django, PHP**, Ruby on Rails, Java Backends
  - JavaScript + AJAX Frontends
- **Databases**
  - MongoDB, MySQL, PostGreSQL, SciDB
  - http://tinyurl.com/nerscdbs
- **Public or Authenticated Gateways**
  - http://portal.nersc.gov OR https://portal-auth.nersc.gov
- **OpenDAP and MQ services**

Use NEWT API to access NERSC

- **NEWT – the NERSC REST API**
- **Use the NEWT HTTP API to access NERSC HPC resources directly from your web apps.**

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"}]
```

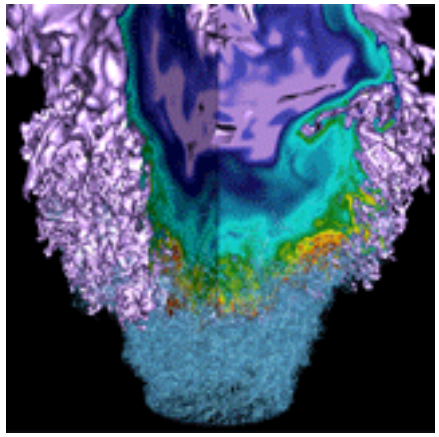
NEWT Features

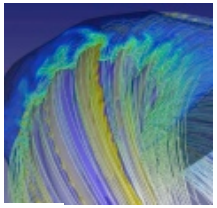
- **Run Commands on any system**
- **Transfer files**
- **Authentication**
- **Submit/Query Jobs directly through NEWT**
- **Persistent Store**
  - Store JSON objects in the NEWT storage
- **Access to NIM info**
  - Information about Users, Repos etc.
- **System Status**

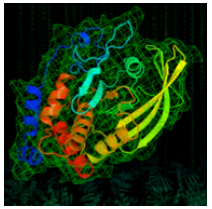
Some Examples

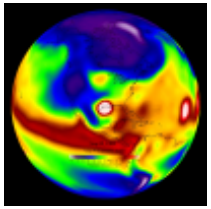
- http://materialsproject.org
- https://spot.nersc.gov
- https://openmsi.nersc.gov
- https://newt.nersc.gov/ for NEWT documentation and live demos

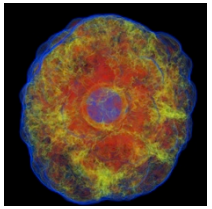
Interactive Environments

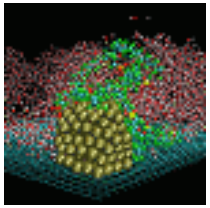






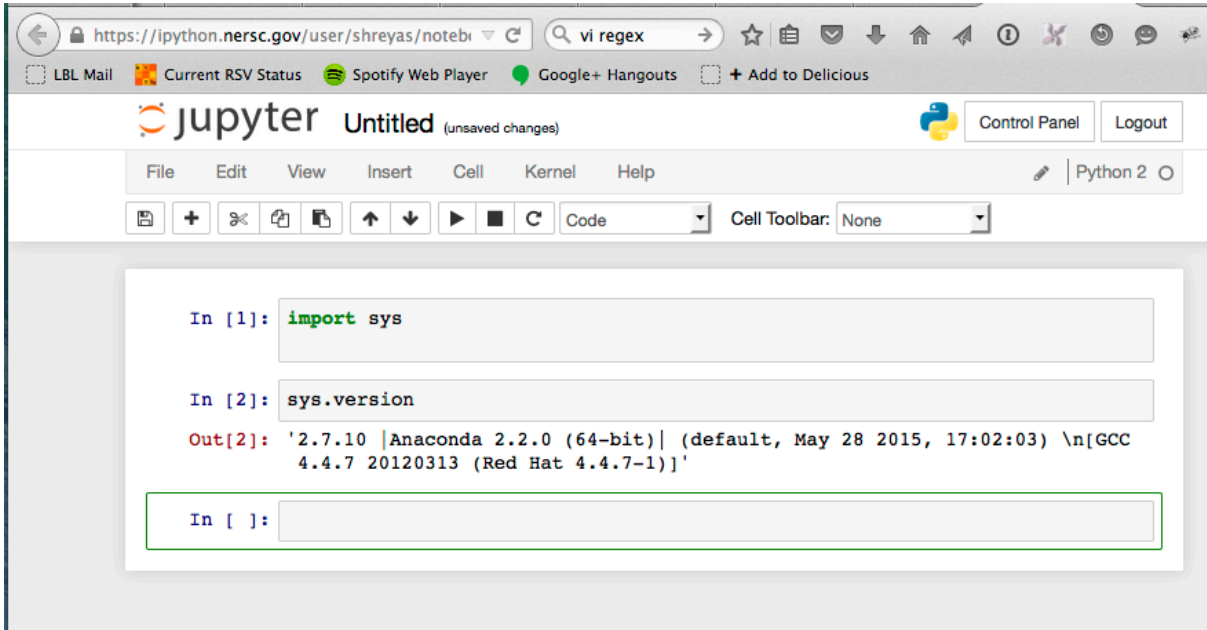






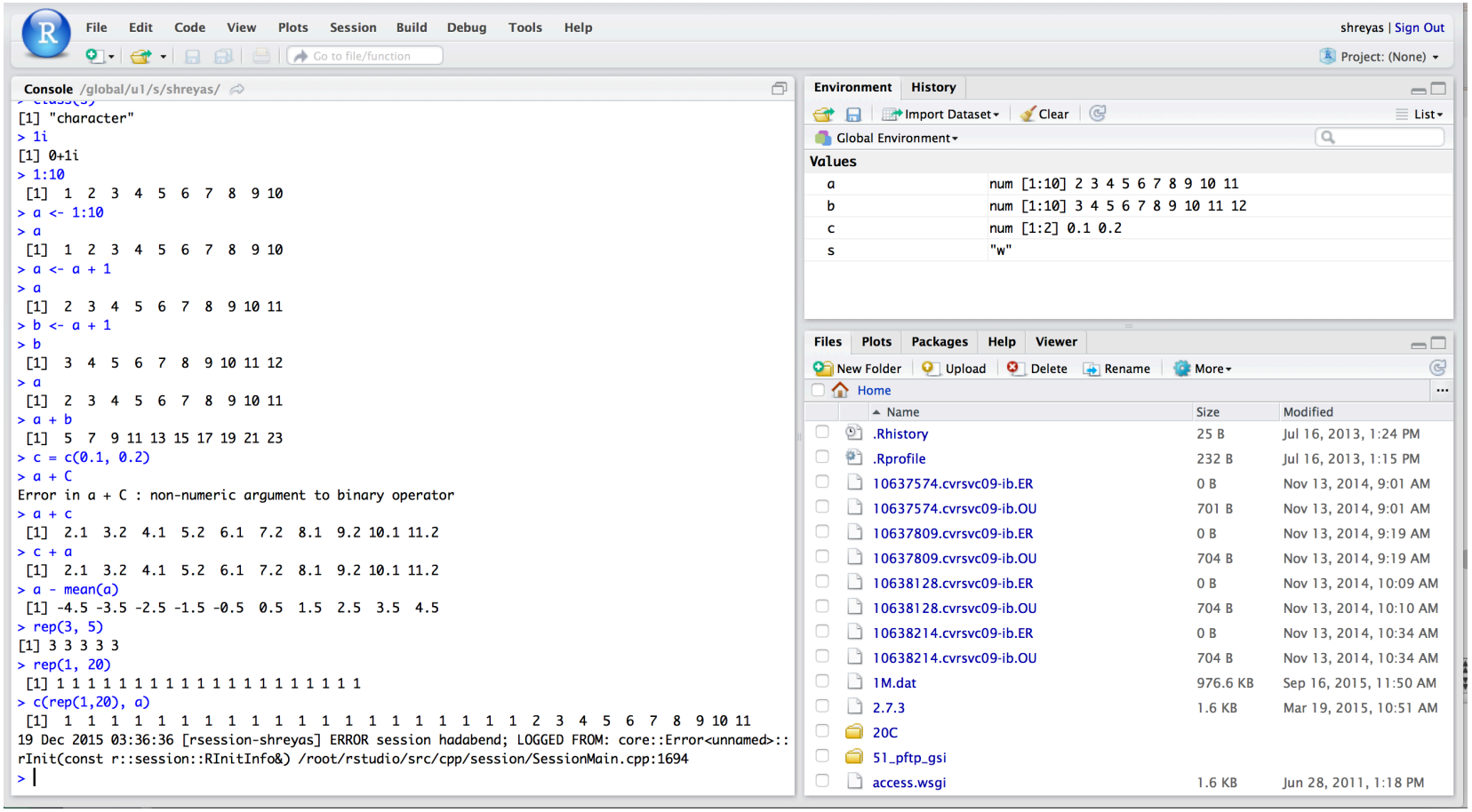
Jupyterhub – iPython Notebook Service

- **Jupyter server with access to NERSC resources**
  - https://ipython.nersc.gov
  - Login with NERSC credentials
- **Access to NERSC global filesystems**
- **Runs inside NERSC network**



R Studio

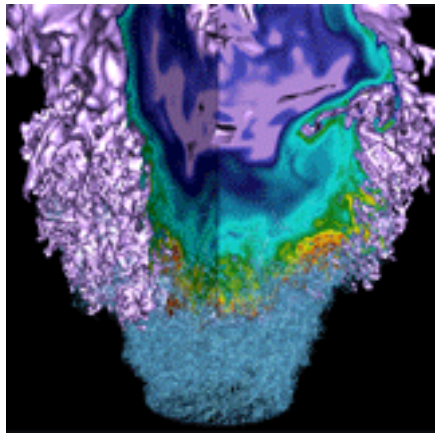
- https://r.nersc.gov

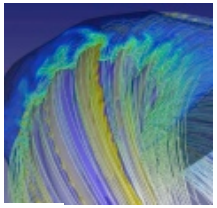


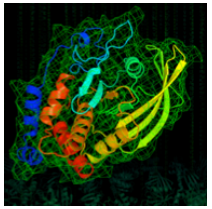
Engagement

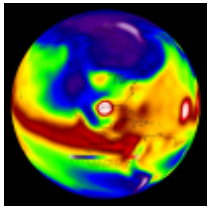
- **There are various levels of engagement and collaboration**
  - Level 1: NERSC provides building blocks and backend infrastructure, science groups build their own gateways.
  - Level 2: NERSC offers some initial consulting help to get you started.
  - Level 3: Immersive development - science groups work in collaboration with NERSC engineers to build gateways. (But requires more resources from both sides).
- **If you are interested in building a portal please come and talk to us. We can help customize our offerings to meet user needs.**
  - Customized VMs, URLs, Docker deployments

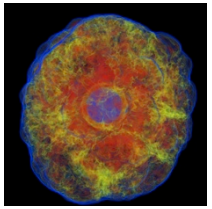
Thanks

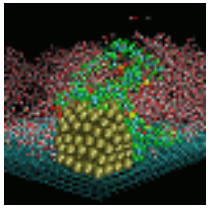






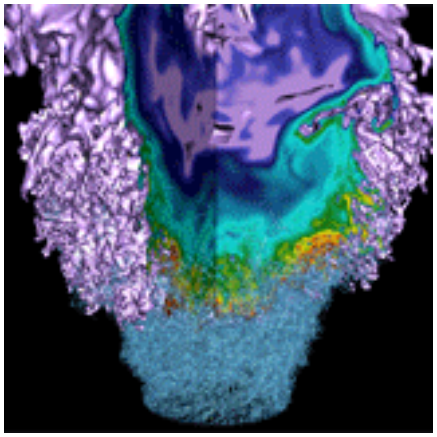


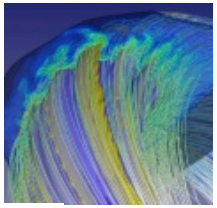


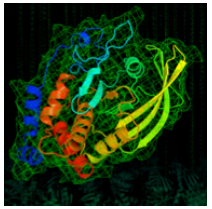


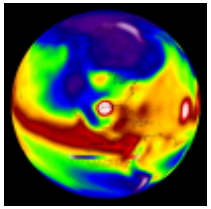
NERSC 40 YEARS at the FOREFRONT

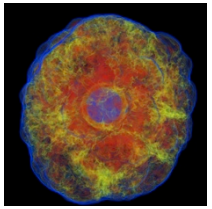
Extra

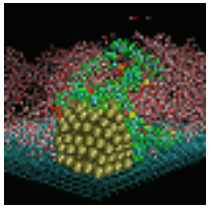












GridFTP

- **Grid tools available on NERSC systems via "module load globus"**
  - globus-url-copy

  http://tinyurl.com/nersc-gridftp

- **Use Data Transfer nodes for wide-area transfers**
  - dtn01.nersc.gov
  - dtn02.nersc.gov
  - dtn03.nersc.gov
  - dtn04.nersc.gov

- **Use "grid" name for system specific file systems (often not optimal solution)**
  - corigrid.nersc.gov
  - edisongrid.nersc.gov
  - pdsfgrid.nersc.gov

BaBar Copy (bbcp)

- **Developed for BaBar experiment at SLAC**
  - Somewhat complicated command-line
    https://www.nersc.gov/users/data-and-file-systems/transferring-data/bbcp/
    - http://tinyurl.com/nerscbbcp

- **Peer-to-peer model (not client-server)**
  - Must be installed on each end
    - Easy to build and/or install
    - Available on all NERSC systems
  - Can do third-party transfers
  - Uses ssh authentication
  - Good for larger files