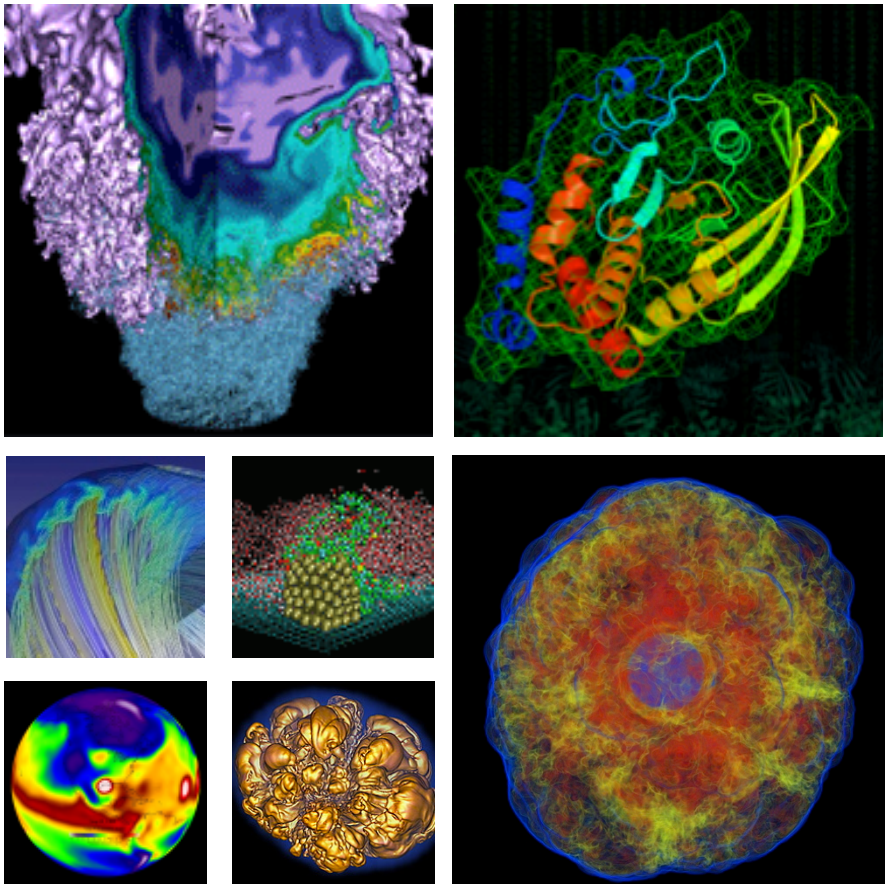
Moving Data at NERSC



Jeff Porter
rjporter@lbl.gov

Data and Analytics Services New User Training 2015

August 13th, 2015

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

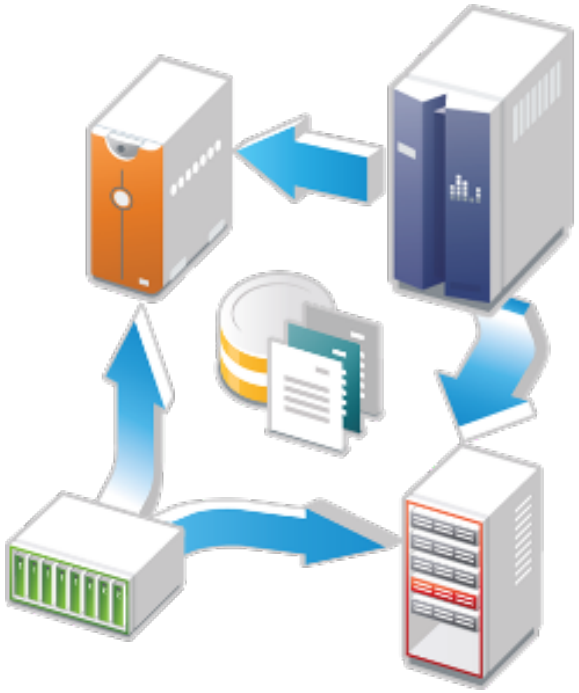
DTN File Systems

- **Global File Systems are available**
  - `/global/homes`
  - `/global/scratch2`
  - `/global/project`
  - `/global/projectb`
  - `/global/projecta`
  - `/global/dna`
  - `/global/seqfs`
  - `/global/common`

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

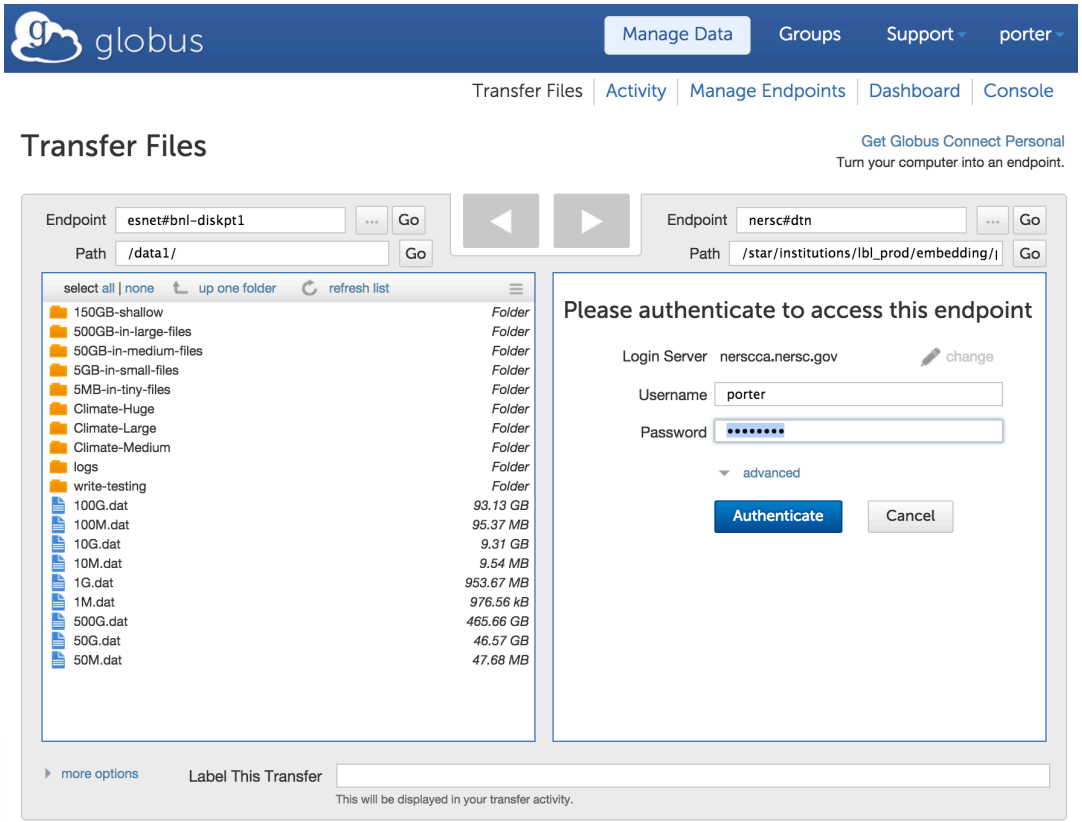
Transfer Tools

- **Globus ~~Online~~ … 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



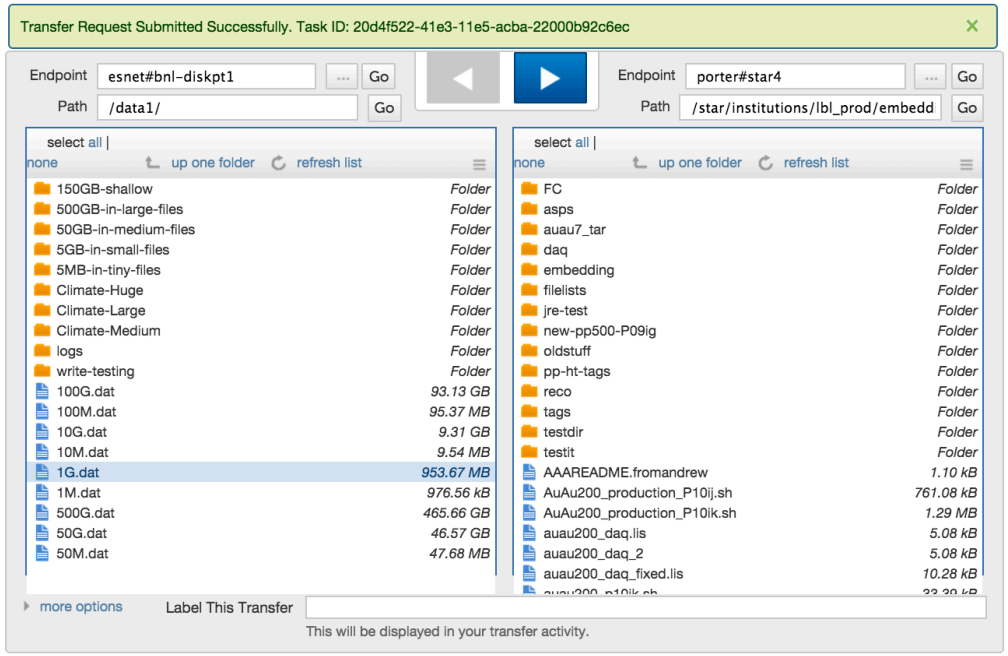
Globus.org

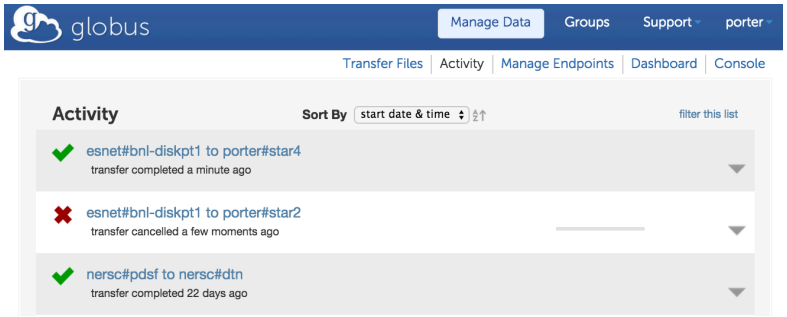
- **Managed 3rd party transfers** https://www.globus.org/
  - Web based interface, currently requires Globus Account
- **Type `nersc#dtn` endpoint in the transfer window**
  - Activate with NERSC username & password



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)

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

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

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**

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)**
  - hoppergrid.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

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/

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

Thank you.