

SciDMZ Data Transfers

(Featuring *Globus* DTNs)
over UCSC's 100GE Network

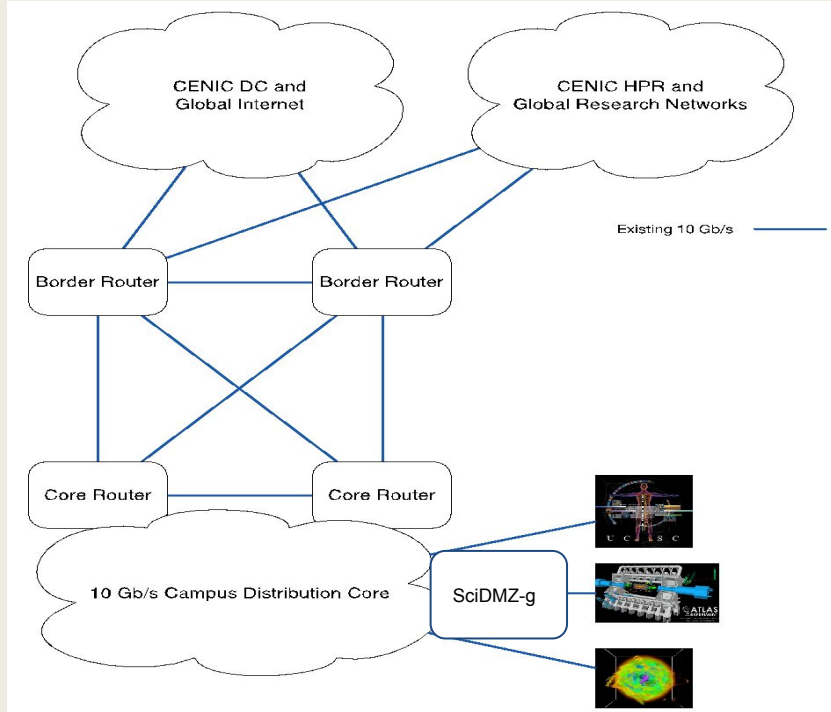
By *Josh Sonstroem* and the **UCSC Hummingbird** team
(jsonstro@ucsc.edu, hummingbird@ucsc.edu)

SciDMZ Data Transfers

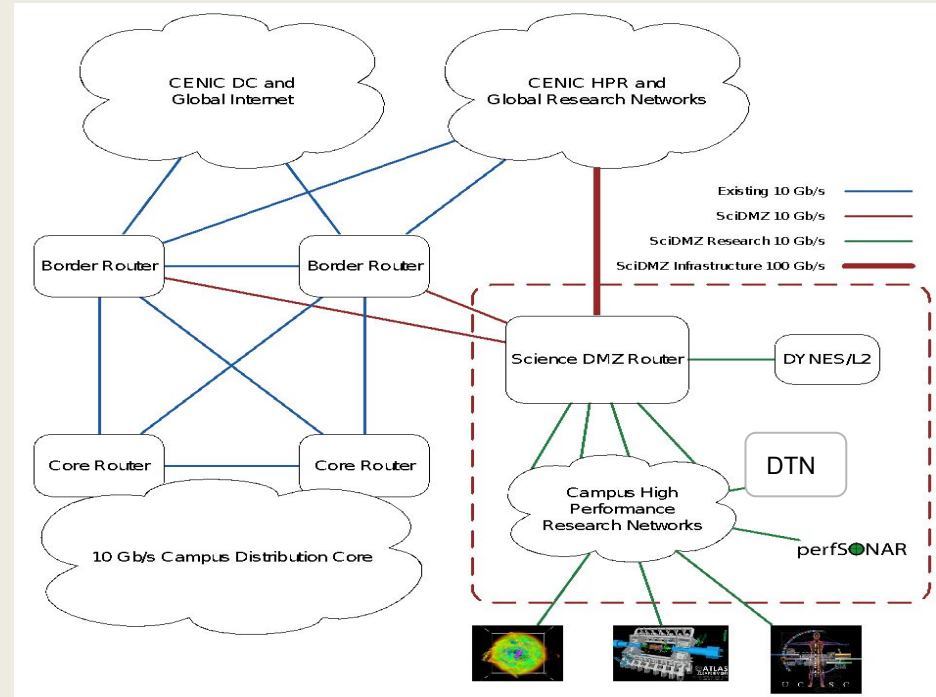
Table of Contents

1. Overview of **UCSC's Science DMZ** (and its clients)
2. Setup and Verify Accounts
 - A. **hbfeeder.ucsc.edu**
 - B. ***Globus Online***
3. Run a Sample ***Globus*** Transfer
4. Install ***Globus Connect Personal*** (or ***Server***)
5. Run a ***PerfSONAR*** Test By-Hand

10GE -> 100GE Science DMZ



OLD (Science traffic flowing thru shared prod borders)



NEW (Science traffic flowing thru 100GE network)

SciDMZ Research Clients

- **Center for Biomolecular Science & Engineering (CBSE)**
 - genomics.ucsc.edu
 - cghub.ucsc.edu – currently 81K files total @ 2PB, downloads 1PB/mo
 - Big Data in Translational Genomics NIH project (BD2K)
- **Santa Cruz Institute for Particle Physics (SCIPP)**
 - ATLAS/LHC Tier-3
 - 10-20TB per year
- **Astrophysics**
 - lux cluster – 3.6PB Lustre Filesystem
 - 100PB+ simulation data at DoE National Labs
- **Campus services**
 - Hummingbird, DTN,
 - PerfSONAR (bwctl10, dps10)

Hummingbird Data Movement

For best performance and to not be flagged for abuse please use hbfeeder.ucsc.edu for ALL data transfers onto and off-of the hb cluster. There are 2 distinct methods available:

- Third-Party Transfer services
 - Globus
 - gridftp
- First-Party Transfer services
 - rsync
 - scp
 - wget/curl

Setup and Verify Accounts (cont.)

3rd Party Transfer -- Authentication and authorization handled outside. Data path directly between client endpoint.

Setup a *Globus Online* Account

1. Visit <http://globus.org>
2. Click the **Sign Up** button at upper right
3. Fill out form with appropriate info
4. Press the **Register** button
5. Finish the process
6. **Sign in** with your *Username/Password*



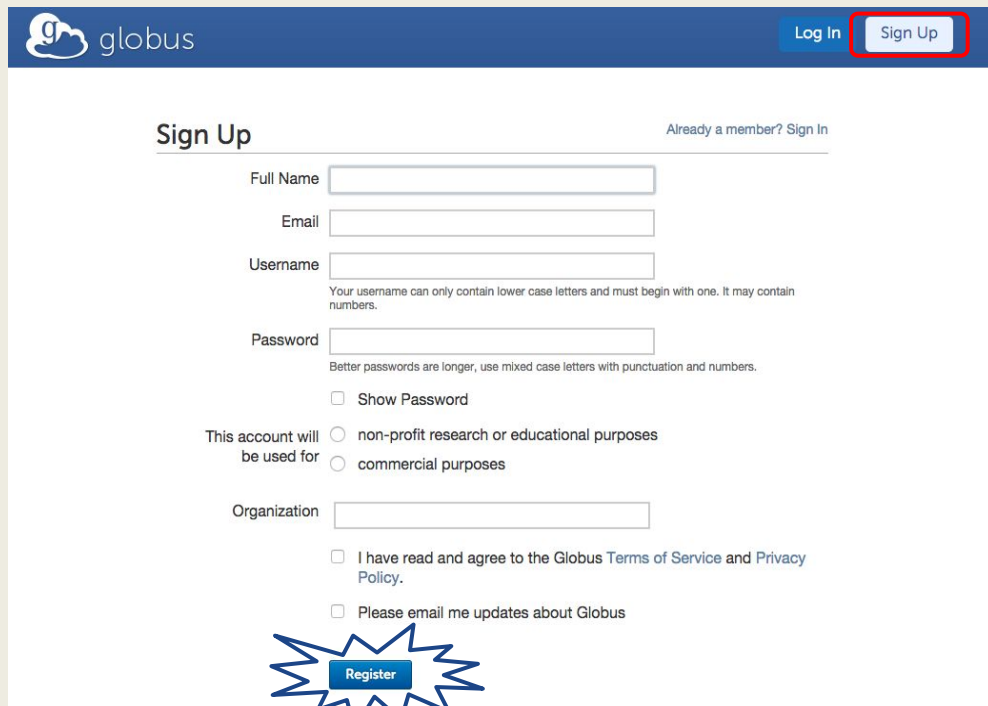
Sign In Sign Up with Globus

Using your Globus login. alternate login

Username

Password

Sign In Forgot password?



globus Log In **Sign Up**

Sign Up Already a member? Sign In

Full Name

Email

Username

Your username can only contain lower case letters and must begin with one. It may contain numbers.

Password

Better passwords are longer, use mixed case letters with punctuation and numbers.

Show Password

This account will be used for non-profit research or educational purposes commercial purposes

Organization

I have read and agree to the Globus [Terms of Service and Privacy Policy](#).

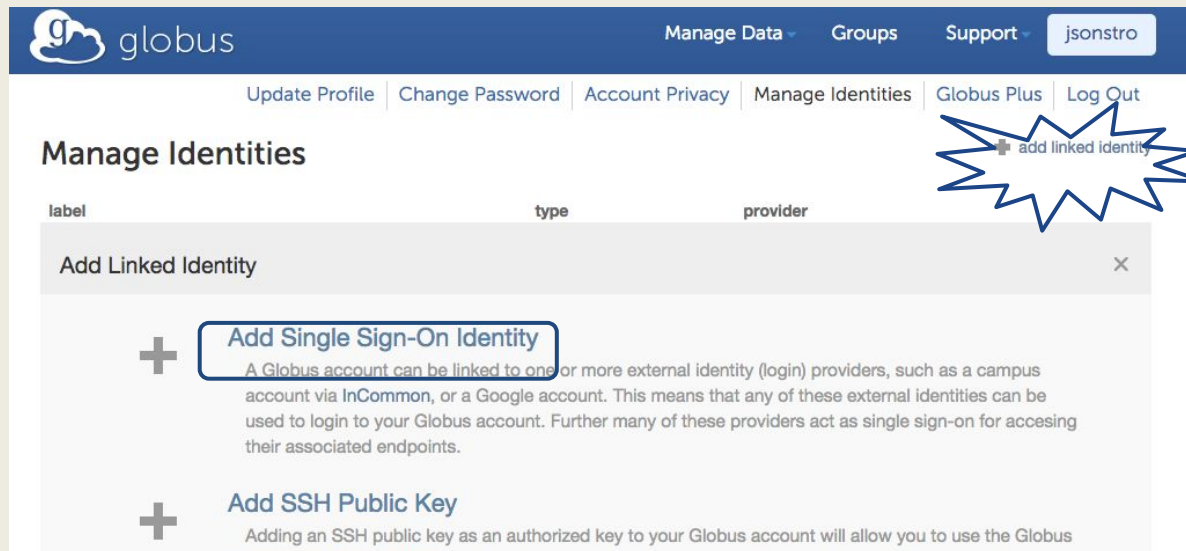
Please email me updates about Globus

Register

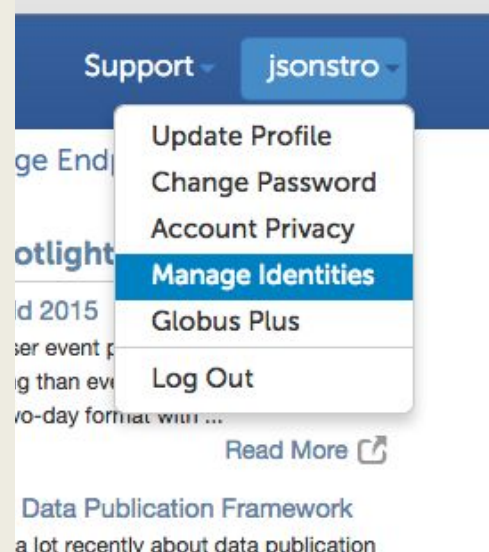
InCommon to *Globus* Mapping

Once logged to your *Globus* account

1. Choose **Manage Identities** from the *drop-down menu*
2. Click **+ add linked identity** at upper right
3. Select the first item, **+ Add Single Sign-on Identity**



The screenshot shows the Globus user interface. At the top, there is a navigation bar with the Globus logo and several menu items: 'Manage Data', 'Groups', 'Support', and 'jsonstro'. Below this, there is a secondary navigation bar with links for 'Update Profile', 'Change Password', 'Account Privacy', 'Manage Identities', 'Globus Plus', and 'Log Out'. The 'Manage Identities' link is highlighted, and a starburst graphic points to a '+ add linked identity' button. A modal window titled 'Add Linked Identity' is open, showing two options: '+ Add Single Sign-On Identity' (circled in blue) and '+ Add SSH Public Key'. The 'Add Single Sign-On Identity' option includes a description: 'A Globus account can be linked to one or more external identity (login) providers, such as a campus account via InCommon, or a Google account. This means that any of these external identities can be used to login to your Globus account. Further many of these providers act as single sign-on for accessing their associated endpoints.'



The screenshot shows a dropdown menu for the 'jsonstro' user. The menu items are: 'Update Profile', 'Change Password', 'Account Privacy', 'Manage Identities' (highlighted in blue), 'Globus Plus', and 'Log Out'. Below the menu, there is a 'Read More' link with an external link icon, and a section titled 'Data Publication Framework' with a sub-heading 'a lot recently about data publication'.

InCommon to *Globus* Mapping (cont.)

- Next, choose *InCommon / CILogon* from the list
- Click the **Proceed** button

label type provider

Add Linked Identity

Add Single Sign-On With: InCommon / CILogon

Label

Click PROCEED to authenticate with InCommon / CILogon.

Proceed Cancel

- Select *University of California, Santa Cruz* and press the **Log On** button

Site Name: Globus
Site URL: https://www.globus.org
Service URL: https://www.globus.org/service/graph/users/jsonstro/credentials/oauth_callback

Select An Identity Provider:

University of Arkansas
University of California, Davis
University of California, San Francisco
University of California, Santa Cruz

Remember this selection:

Log On

By selecting "Log On", you agree to CILogon's privacy policy.

Manage Identities

label	type	provider
Add Linked Identity		
Select Identity Provider		
Select the single sign-on provider you wish to associate with your Globus account		
Argonne LCF		LRZ
Argonne MCS & LCRC		NCAR RDAC
BIRN		NCSA
CLI Transition		NCSA Blue Waters
EGIG		NERSC
ESG ANL		Tuakiri
Exeter		UChicago CI
Google		UChicago iBi
InCommon / CILogon		WestGrid
		XSEDE

- Finally, authenticate with your *CruzID* and your **Gold** password

UC SANTA CRUZ

CruzID

CruzID Gold Password

Login

CruzID Gold



Run a Sample *Globus* Transfer

Transfer Files | Activity | Manage Endpoints | Dashboard

Get Globus Connect Personal
Turn your computer into an endpoint.

Endpoint ... Go

Path Go

Endpoint ... Go

Path

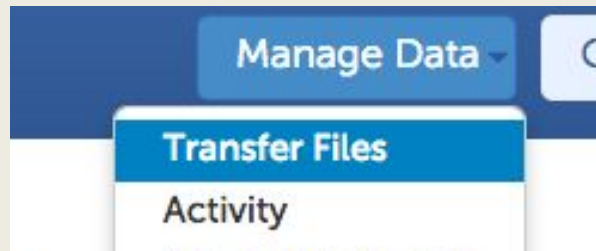
select all | none | up one folder | refresh list

- brad Folder
- hauskins Folder
- jackman Folder
- jsonstro Folder
- rdreece Folder
- shaw Folder

Please select an endpoint above.

more options Label This Transfer

This will be displayed in your transfer activity.



1. Choose **Transfer Files** from menu bar or the **Manage Data** dropdown menu at top
2. Enter **ucsc#hb_home** in one of the two *Endpoint* fields then enter your homedir
3. Enter **esnet#** in the other *Endpoint* field and choose **ESnet Read-Only Test at LBL DEV**
4. Click into your **CruzID** directory

Run a Sample *Globus* Transfer (cont.)

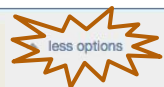


Endpoint: ucsc#dtn Path: /data/ Go

Endpoint: esnet#bni-diskpt1 Path: /data1/ Go

Folder	Folder
brad	500GB-in-large-files
hauskins	50GB-in-medium-files
jackman	5GB-in-small-files
jsonstro	5MB-in-tiny-files
rdreece	write-testing
shaw	100G.dat
	100M.dat
	10G.dat
	10M.dat
	1G.dat
	1M.dat
	500G.dat
	50G.dat
	50M.dat

5. Click on a *file* or *directory* in the **esnet#** endpoint
6. Press the highlighted **blue arrow** at center to begin the transfer
7. Watch it go and click the **arrow** under **Activity** to see more info
8. Select **More Options** for **Transfer Settings** like *rsync*-style add-ons, encryption, preserve mod times, send deltas, etc.



Label This Transfer

This will be displayed in your transfer activity.

Transfer Settings

- only transfer new or changed files ?
- delete files on destination that do not exist on source ?
- preserve source file modification times ?
- verify file integrity after transfer ?
- encrypt transfer ?

Run a Sample *Globus* Transfer (cont.)

9. Watch the **Activity** window for updates about the transfer's progress
10. Click **View Debug Data** in the bottom right for a more holistic view

The screenshot shows the 'Activity' window for a Globus transfer. The window title is 'esnet#bnl-diskpt1 to ucsc#dtn' and it indicates the transfer started a minute ago. There are two tabs: 'Overview' (selected) and 'Event Log'. A 'cancel task' button is visible in the top right of the overview panel. The transfer details are as follows:

Task ID	184c2cd6-f2a6-11e4-ab4a-22000b92c6ec
Source	esnet#bnl-diskpt1
Destination	ucsc#dtn
Condition	ACTIVE
User	jsonstro
Requested	2015-05-04 02:39 pm
Deadline	2015-05-05 02:39 pm
Transfer Settings	<ul style="list-style-type: none">• overwriting all files on destination• transfer is not encrypted

Summary statistics:

Files	1
Directories	0
Bytes Transferred	30.13 GB
Effective Speed	3.39 Gbit/s
Pending	1
Succeeded	0
Cancelled	0
Expired	0
Failed	0
Retrying	0
Skipped	0

A blue starburst icon labeled 'view debug data' is located at the bottom right of the overview panel.

The screenshot shows the 'Debug Data' window, which provides a detailed log of the transfer. A 'copy to clipboard' button is in the top right corner. The data is as follows:

```
display_bytes_transferred 32348045312
display_effective_speed 334864153
display_bytes_pretty 30.13 GB
display_mbps 2676.91
display_label esnet#bnl-diskpt1 to ucsc#dtn
display_source_endpoint esnet#bnl-diskpt1
display_dest_endpoint ucsc#dtn
refined_status active
display_refined_status transfer_active
display_time_since 2 minutes ago
display_duration 00:01:36
is_transfer true
is_delete false
task_id 184c2cd6-f2a6-11e4-ab4a-22000b92c6ec
username jsonstro
bytes_transferred 32348045312
DATA_TYPE task
deadline 2015-05-05 21:39:52+00:00
type TRANSFER
destination_endpoint ucsc#dtn
effective_bytes_per_second 334864153
files 1
delete_destination_extra false
request_time 2015-05-04 21:39:52+00:00
nice_status OK
subtasks_expired 0
subtasks_cancelled 0
faults 0
subtasks_total 1
nice_status_expires_in 863692
subtask_link [object Object]
status ACTIVE
bytes_checksummed 0
```

Run a Sample *Globus* Transfer (cont.)

Here is an example of a completed transfer's **Activity** window and entire **Debug Data** screen at right. Average speed was:

- **5.16Gb/s (100Gb)** from **Brookhaven (NL), NY** to **UCSC**

Activity

✔ **esnet#bnl-diskpt1 to ucsc#dtn**
transfer completed a few moments ago

Overview | Event Log

Task ID	184c2cd6-f2a6-11e4-ab4a-22000b92c6ec
Source	esnet#bnl-diskpt1 ~ 100Gb
Destination	ucsc#dtn
Condition	SUCCEEDED
User	jsonstro
Requested	2015-05-04 02:39 pm
Completed	2015-05-04 02:42 pm

transfer in just 3 min!

Files	1
Directories	0
Bytes Transferred	93.13 GB
Effective Speed	5.16 Gbit/s
Pending	1
Succeeded	1
Cancelled	0
Expired	0
Failed	0
Retrying	0
Skipped	0

Transfer Settings

- overwriting all files on destination
- transfer is not encrypted

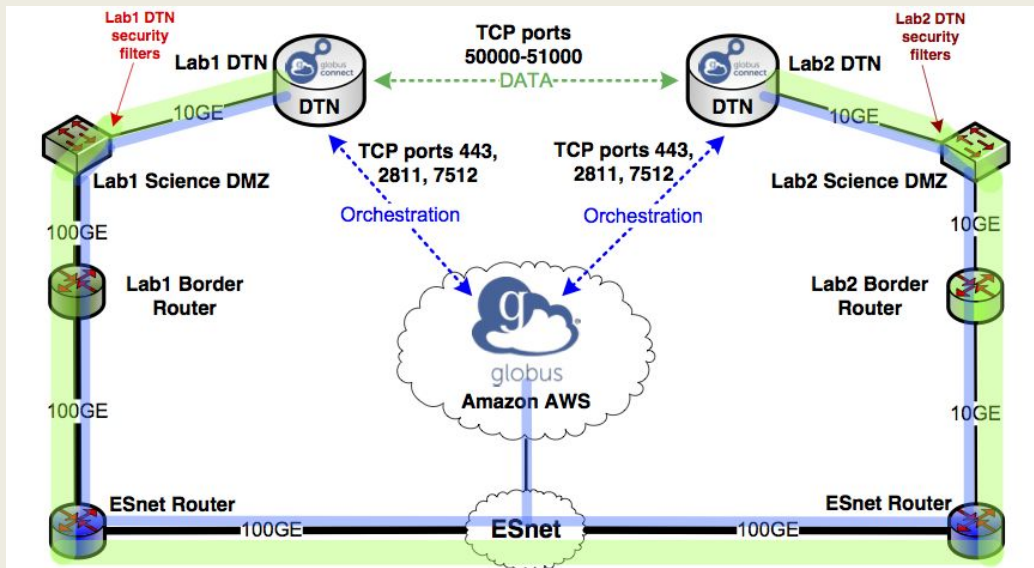
hide debug data

```
Debug Data
copy to clipboard

display_bytes_transferred 10000000000
display_effective_speed 645230705
display_bytes_pretty 93.13 GB
display_abps 5161.85
display_label esnet#bnl-diskpt1 to ucsc#dtn
display_source_endpoint esnet#bnl-diskpt1
display_dest_endpoint ucsc#dtn
refined_status success
display_refined_status transfer_success
display_time_since 3 minutes ago
display_duration 00:02:35
is_transfer true
is_delete false
task_id 184c2cd6-f2a6-11e4-ab4a-22000b92c6ec
username jsonstro
bytes_transferred 10000000000
DATA_TYPE task
completion_time 2015-05-04 21:42:27+00:00
deadline 2015-05-05 21:39:52+00:00
type TRANSFER
destination_endpoint ucsc#dtn
effective_bytes_per_second 645230705
files 1
delete_destination_extra false
request_time 2015-05-04 21:39:52+00:00
subtasks_expired 0
subtasks_canceled 0
faults 0
subtasks_total 1
subtask_link [object Object]
status SUCCEEDED
bytes_checksummed 0
subtasks_failed 0
history_deleted false
files_skipped 0
subtasks_retrying 0
preserve_timestamp false
event_link [object Object]
encrypt_data false
source_endpoint esnet#bnl-diskpt1
subtasks_succeeded 1
command API 0.10 go
subtasks_pending 1
verify_checksum false
directories 0
```

Globus Transfer “Under the Hood”

```
$ globus-url-copy -tcp-bs 12M -bs 12M -p 8 -fast gsiftp://user@dtm.snln.gov/lustre/user/10G.dat gsiftp://user@dtm.ucsc.edu/data/user/10G.dat
```



Logical data path

Physical data path

Logical control path

Physical control path

Lab1 DTN security filters

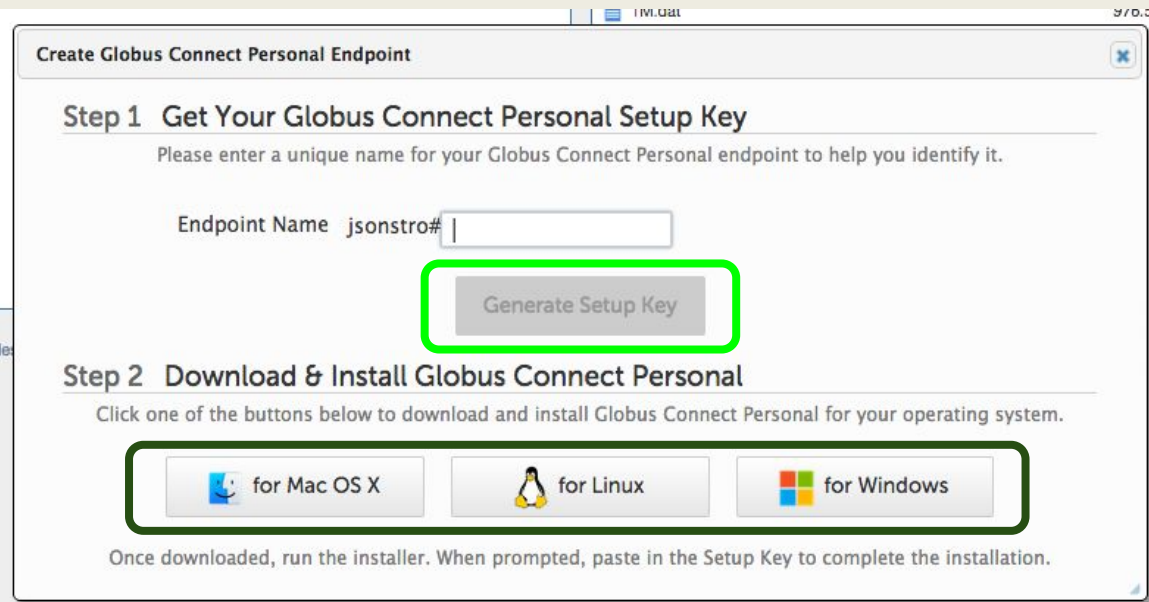
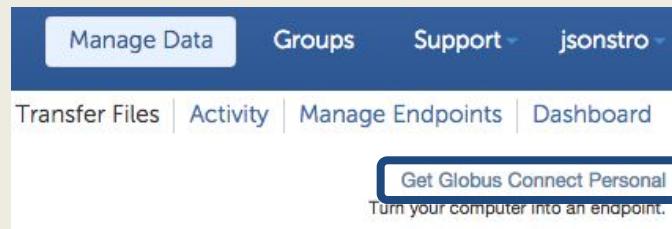
Src Address	Src Port	Dst Address	Dst Port
Lab1 DTN	TCP 50000-51000	Lab2 DTN	TCP 50000-51000
Lab1 DTN	TCP 443, 2811, 7512	Globus Cloud	TCP unprivileged
Lab2 DTN	TCP 50000-51000	Lab1 DTN	TCP 50000-51000
Globus Cloud	TCP unprivileged	Lab1 DTN	TCP 443, 2811, 7512

Lab2 DTN security filters

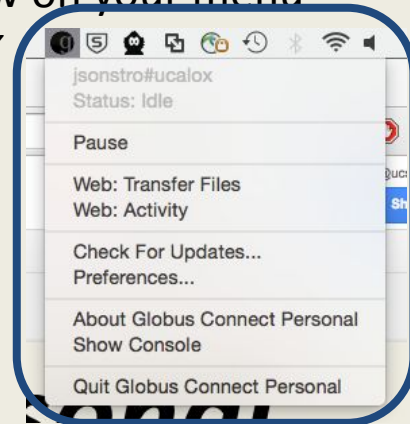
Src Address	Src Port	Dst Address	Dst Port
Lab2 DTN	TCP 50000-51000	Lab1 DTN	TCP 50000-51000
Lab2 DTN	TCP 443, 2811, 7512	Globus Cloud	TCP unprivileged
Lab1 DTN	TCP 50000-51000	Lab2 DTN	TCP 50000-51000
Globus Cloud	TCP unprivileged	Lab2 DTN	TCP 443, 2811, 7512

Install *Globus Connect Personal*

1. On the *Transfer Files* screen, click **Get Globus Connect Personal** at upper right
2. Enter an *Endpoint Name* for your laptop
3. Click the **Generate Setup Key** button
4. Then *download* and *install* the software for your OS



5. Once installed and running it will show on your menu bar in *OSX*



Install *Globus Connect Personal* (cont.)

6. Your laptop is now visible to your user (and only your user) via *Globus* (both web and CLI)

The screenshot shows the Globus web interface. At the top, there's a navigation bar with the Globus logo, 'Manage Data', 'Groups', 'Support', and 'jsonstro'. Below this, there are tabs for 'Transfer Files', 'Activity', 'Manage Endpoints', and 'Dashboard'. The main heading is 'Transfer Files'. On the right, there's a link to 'Get Globus Connect Personal' and the text 'Turn your computer into an endpoint.' The interface is split into two panels. The left panel shows the local endpoint 'jsonstro#ucalox' with the path '/~/'. The right panel shows the remote endpoint 'ucsc#dtn' with the path '/data/jsonstro/'. Both panels have a file list. The left panel shows a list of folders like Applications, Desktop, Documents, etc. The right panel shows a list of files: 100G.dat (93.13 GB), 10G.dat (9.31 GB), 1G.dat (947.00 MB), and 50G.dat (46.57 GB). At the bottom, there's a 'Label This Transfer' field and a note 'This will be displayed in your transfer activity.'

CONGRATS!!!

Now you can use
your laptop to
upload/download
to/from any other
Globus node

*Just remember it limits your
max-speed to the rate of your
laptop network connection*



Run a BWCTL Test By-Hand

```
% bwctl -f m -T iperf3 -i 2 -s perf-scidmz-data.cac.washington.edu -c dps10.ucsc.edu -t 30 -w 128M
```

```
SENDER START
Connecting to host 128.114.109.66, port 5257
[ 14] local 198.124.238.146 port 55387 connected to 128.114.109.66 port 5257
[ ID] Interval          Transfer      Bandwidth    Retr  Cwnd
[ 14]  0.00-2.00      sec    335 MBytes   1405 Mbits/sec    0   2.99 MBytes
[ 14]  2.00-4.00      sec   72.5 MBytes    304 Mbits/sec    0   7.58 MBytes
[ 14]  4.00-6.00      sec   165 MBytes    692 Mbits/sec    0  15.3 MBytes
[ 14]  6.00-8.00      sec   374 MBytes   1568 Mbits/sec    0  43.6 MBytes
[ 14]  8.00-10.00     sec   1.11 GBytes   4771 Mbits/sec    0  128 MBytes
[ 14] 10.00-12.00     sec   2.16 GBytes   9280 Mbits/sec    0  179 MBytes
[ 14] 12.00-14.00     sec   2.15 GBytes   9233 Mbits/sec    0  187 MBytes
[ 14] 14.00-16.00     sec   2.08 GBytes   8944 Mbits/sec    0  187 MBytes
[ 14] 16.00-18.00     sec   2.08 GBytes   8934 Mbits/sec    0  187 MBytes
[ 14] 18.00-20.00     sec   2.08 GBytes   8939 Mbits/sec    0  187 MBytes
[ 14] 20.00-22.00     sec   2.07 GBytes   8881 Mbits/sec    0  187 MBytes
[ 14] 22.00-24.00     sec   2.08 GBytes   8913 Mbits/sec    0  187 MBytes
[ 14] 24.00-26.00     sec   2.07 GBytes   8876 Mbits/sec    0  187 MBytes
[ 14] 26.00-28.00     sec   2.14 GBytes   9191 Mbits/sec    0  187 MBytes
[ 14] 28.00-30.00     sec   2.07 GBytes   8902 Mbits/sec    0  187 MBytes
-----
[ ID] Interval          Transfer      Bandwidth    Retr
[ 14]  0.00-30.00     sec   23.0 GBytes  6589 Mbits/sec    0
[ 14]  0.00-30.00     sec   22.8 GBytes  6536 Mbits/sec
iperf Done.
```

BWCTL Test
London to UCSC:
9.1Gb/s peak
Average (30s):
6.5Gb/s

SciDMZ Data Transfers

(Featuring *Globus* DTNs)

ESnet PerfSONAR Endpoints

anl-pt1 lbl-pt1 lond-pt1

bnl-pt1 amst-pt1

ESnet Globus Test Endpoints

anl-diskpt1 lbl-diskpt1 lond-diskpt1

bnl-diskpt1 amst-diskpt1

SciDMZ Data Transfers

(Featuring *Globus* DTNs)

Additional questions?

Need more help?

Thanks for your time!

High-bandwidth Protocol Testbed

For discovery/analysis the CCNIE team has also deployed a *dummysnet* testbed to simulate different real-world scenarios



- Dummysnet lets us specify path characteristics
 - b/w limits, loss rates, etc.
- Understand impact of configuration parameters on performance
- Determine optimal configurations
- Compare performance