Data Management and Transfer Tools available at SciNet

Ramses van Zon

SciNet HPC Consortium

February 11, 2015

Introduction

Who, What, When, Where, Why, How?

1. Someone has data.	← Who?
2. Data is somewhere.	← Where?
3. You made it or you have to get it.	← How?
4. You analyze it or store it.	← What/Why?
5. You can then delete the data.	← When?

The last step offends many. We'll discuss it at the end.

SciNet compute • calcul CANADA

Outline

1. Storage Places
2. Data (re)organization
3. Moving Data
4. Data plan

Storage Places

Storage Places

- $HOME
- $SCRATCH
- $PROJECT
- $ARCHIVE
- Ram Disk
- Other consortia/clusters
- Internet
- Your own/lab computer

Storage Places

- $HOME
- $SCRATCH
- $PROJECT
- $ARCHIVE
- Ram Disk
- Other consortia/clusters
- Internet
- Your own/lab computer

---

10 GB, backed up, on parallel file system with spinning disks.

Shared among nodes.

Read-only in compute jobs.

Storage Places

- $HOME
- $SCRATCH
- $PROJECT
- $ARCHIVE
- Ram Disk
- Other consortia/clusters
- Internet
- Your own/lab computer

Up to 20TB or 1 million files.

Not backed up, on parallel file system with spinning disks.

Shared among nodes.

Writable from every node.

Purged after three months.

Storage Places

- $HOME
- $SCRATCH
- $PROJECT
- $ARCHIVE
- Ram Disk
- Other consortia/clusters
- Internet
- Your own/lab computer

Up to 5TB, by allocation

Not backed up, on parallel

file system with spinning disks.

Shared among nodes.

Writable from every node.

Not purged.

Storage Places

- $HOME
- $SCRATCH
- $PROJECT
- $ARCHIVE
- Ram Disk
- Other consortia/clusters
- Internet
- Your own/lab computer

2TB/group, more by allocation.

Part disk, part Tape backed.

Complements $PROJECT.

Access through HPSS queue.

Not mounted on login, dev or compute

Not backed up.

Not purged.

Storage Places

- $HOME
- $SCRATCH
- $PROJECT
- $ARCHIVE
- Ram Disk
- Other consortia/clusters
- Internet
- Your own/lab computer

Up to 12 GB. Located in RAM

Access only per node.

Not backed up.

Purged after each job.

Storage Places

- $HOME
- $SCRATCH
- $PROJECT
- $ARCHIVE
- Ram Disk
- Other consortia/clusters
- Internet
- Your own/lab computer

Size depends.

Access over internet.

No internet on compute nodes.

Use login, devel or datamover nodes

ssh, scp, wget, Globus

Data Organization

Linux Commands

We (should) all know the commands to move and copy data on a node:

```
cp x y          # copy file x to file y
mv x y          # move file x to file y, much faster
cp -r a b       # recursively copy folder a into folder b
rsync avz a b # copy not-yet copied content of into folder b
man rsync       # read the man pages!
```

You should also know the commands to compress and pack data into zip or tarballs.

```
tar cf y.tar a      # pack folder a into tarball y.tar
tar caf y.tar.gz a # pack+compress folder a into tarball y.tar.gz
zip y.zip a         # pack+compress folder a into zipfile y.zip
man tar
man zip
```

Data Organization, SciNet Tools: diskUsage

To know how much space you're taking:

```
module load extras
diskUsage
```

More on diskUsage:

```
diskUsage -h
```

or http://wiki.scinethpc.ca/wiki/index.php/Data_Management

Be careful using `du`, as it can be quite slow, especially on large number of files (which you don't have, right?).

Data Organization, SciNet Tools: ish

To know what's in a tar ball:

```
tar taf y.tar.gz      # long listing of tarval content
ish index y.tar.gz    # create index file
ish y.tar.gz.igz      # ls, du, cd inside this index file
```

More on ish:

```
ish help
```

or http://wiki.scinethpc.ca/wiki/index.php/ISH

Moving Data

The Old Way

- To the tape backed storage: HPSS (later)
- From your home computer to SciNet (`$HOME` or `$SCRATCH`), or between Compute Canada sites.

```
scp this USER@login.scinet.utoronto.ca:PATHTODESTINATION
man scp
```

- This works well below about 10GB, and is still okay!

Larger transfers we would used to tell people to use our datamover1 node, which has a faster connection to the outside and does not time out. You would have to initiate the transfer from that datamover1 node.

There now is a faster, more convenient, and more fault-tolerant way...

Globus

Globus

- Data transfer service
- Uses Globus Toolkit (GridFTP) to enable high performance data transfer
- Offers a simple graphical interface
- Easy to use: 'fire and forget'
- Rsync compatibility
- There's a command-line interface too (`module load globus`)

Globus: Getting started

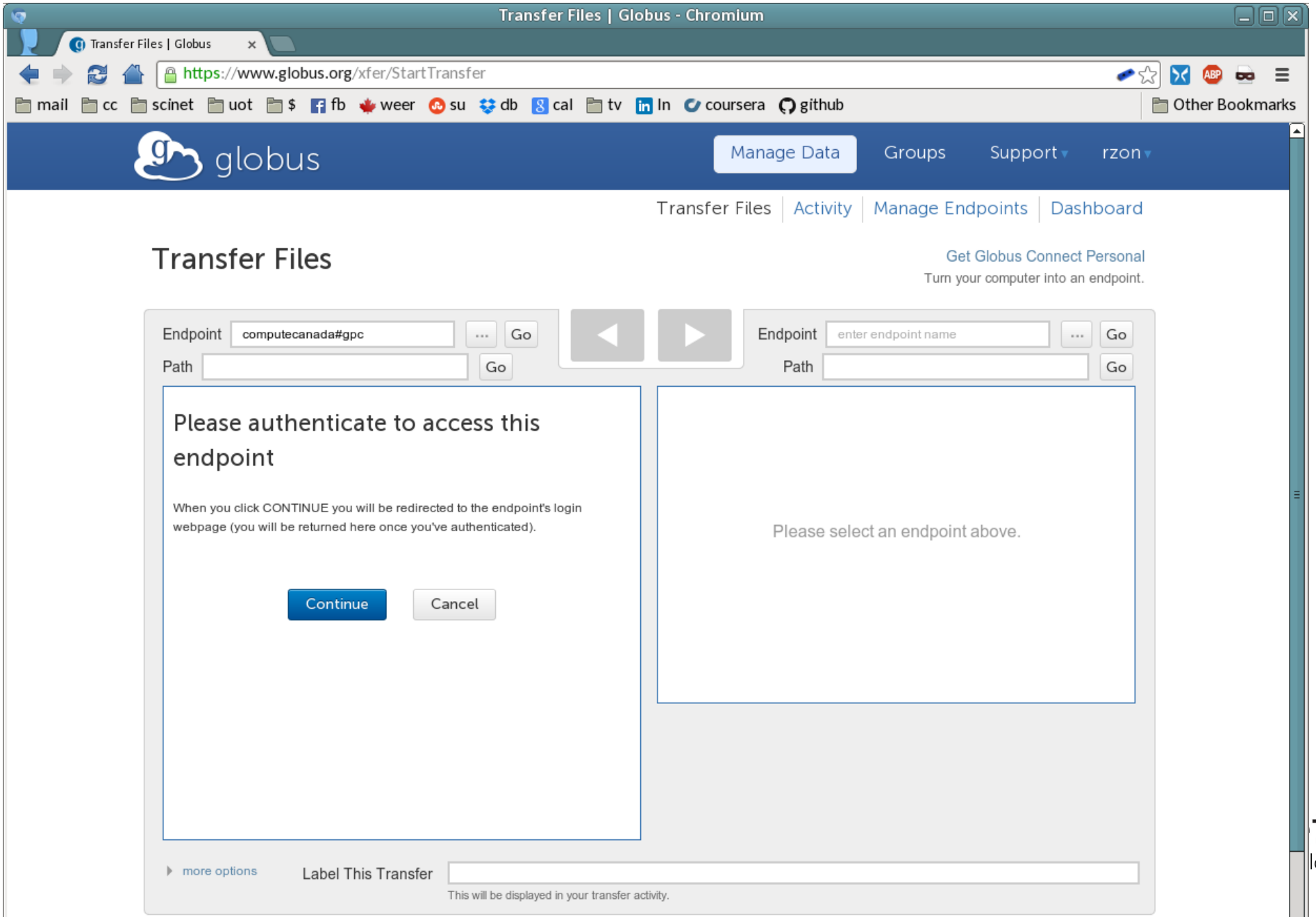
- Setup an account at https://www.globus.org
- Click on "Sign up"
- Fill out the form
- Wait for confirmation

You can now transfer files from the various compute canada site (End-Points) that you have access too, through the same globus.org site (Quick Links -> Transfer Files).

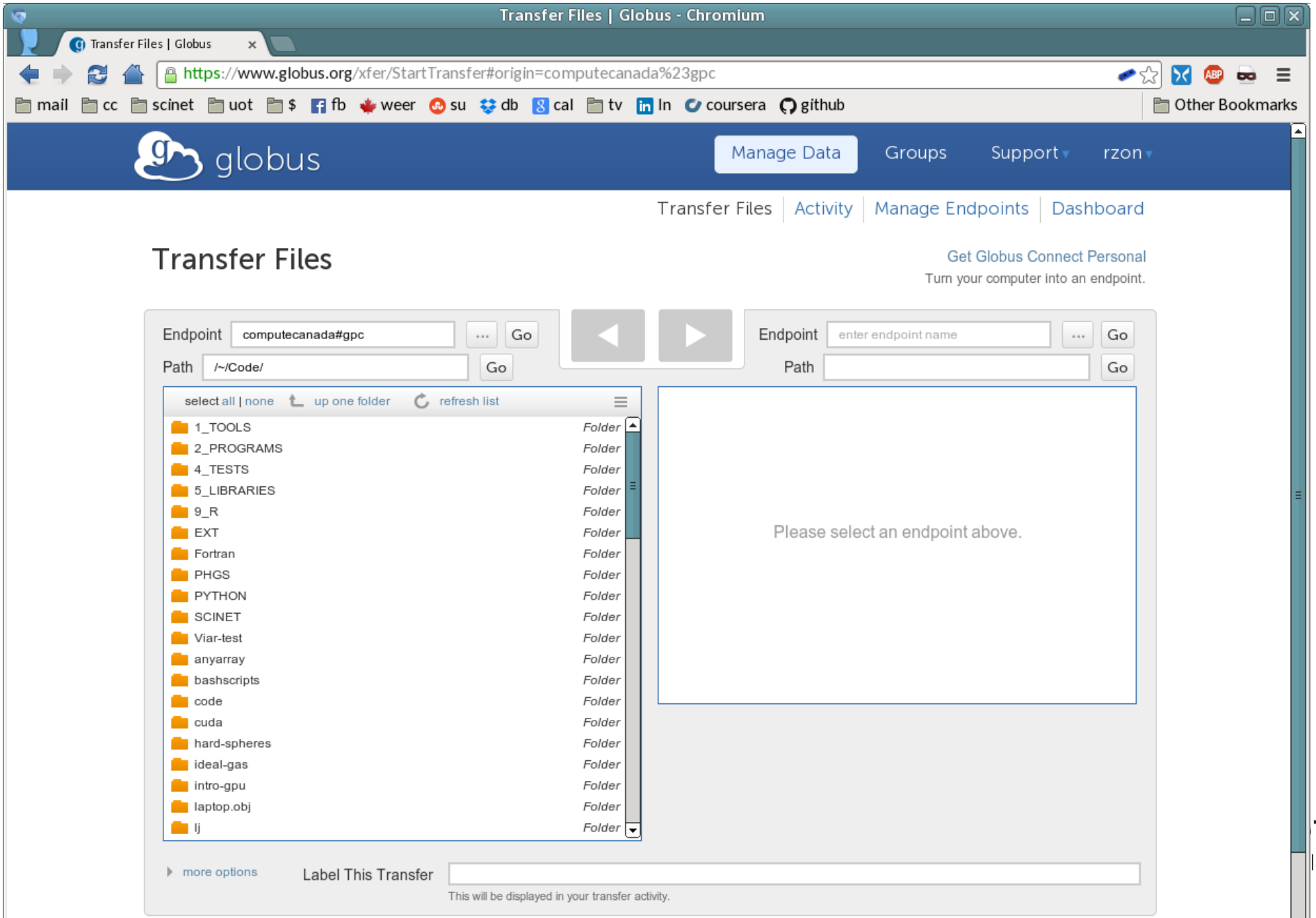
E.g. `computecanada#gpc` is the end-point for SciNet's `$HOME` and `$SCRATCH`.

For each End-Point you select, you will have to log in with your password for that site.

Screenshot 1



Screenshot 2



Globus: Transfer from your own machine

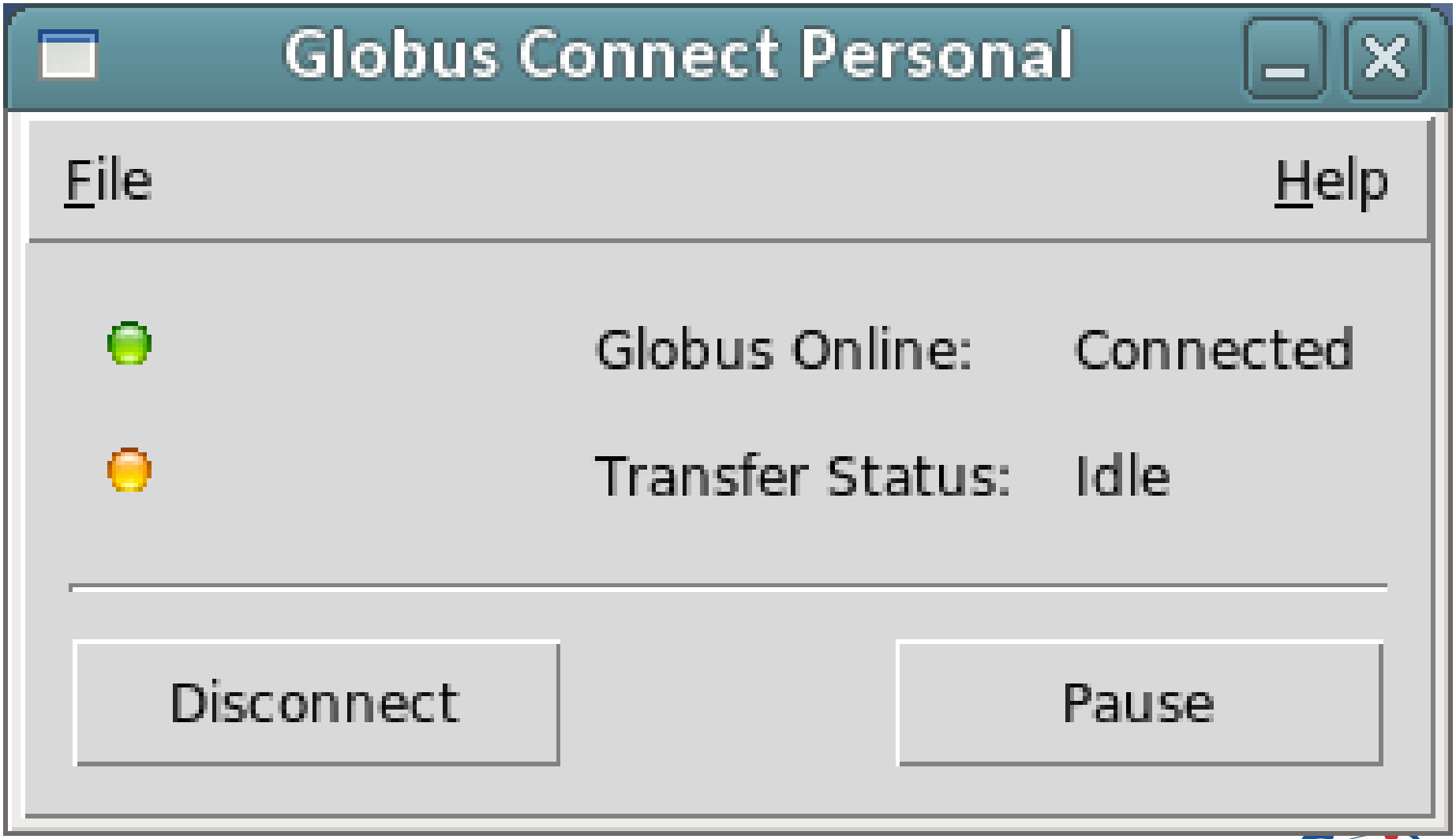
- For this, you need to turn your computer into an End-Point.
- This is done by a program to run on your computer called Globus Connect Personal.
- Download at

https://www.globusonline.org/xfer/ManageEndpoints?globus_connect=true

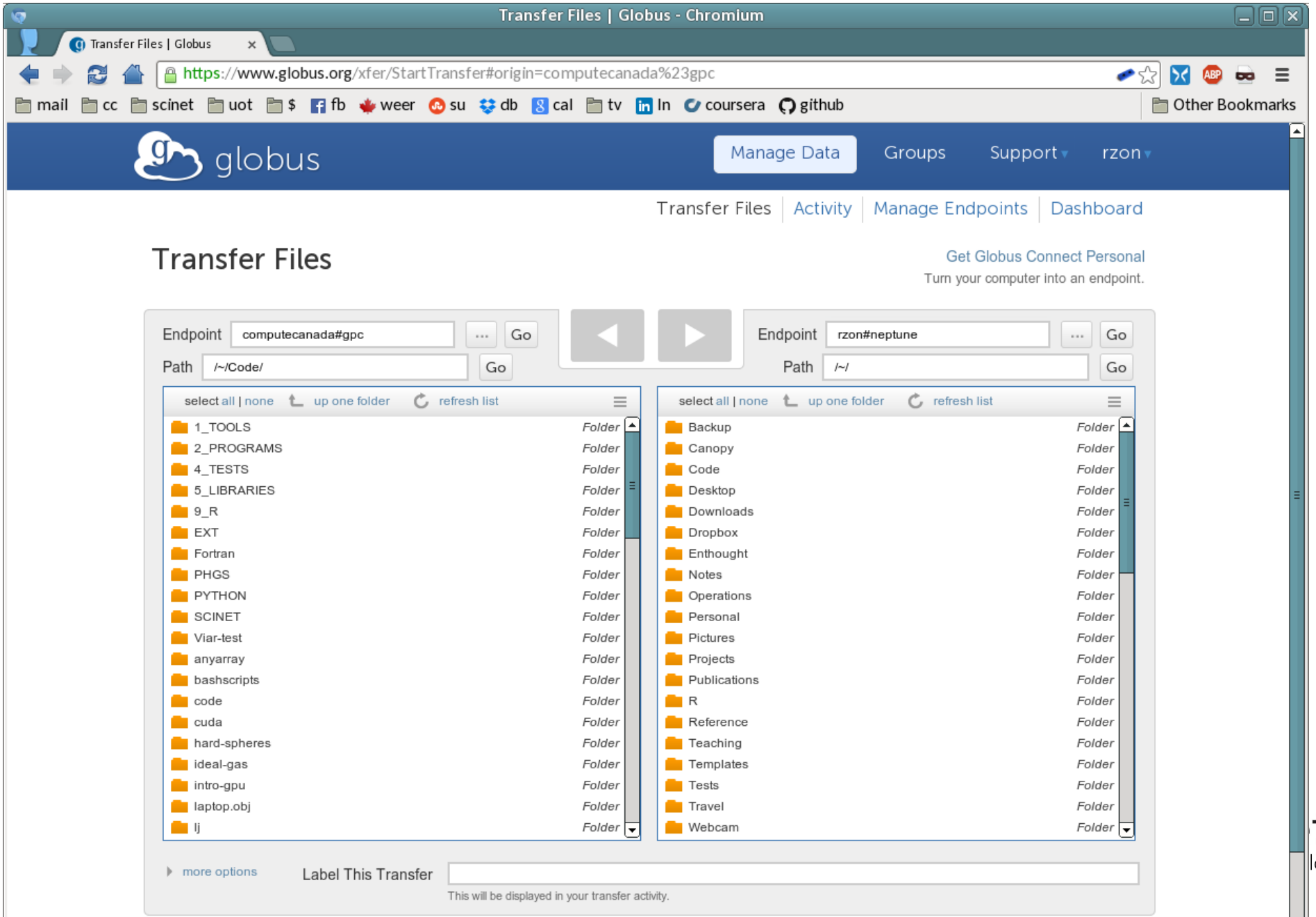
- See following on instruction to setup for your OS on your computer:
  Windows: http://tinyurl.com/p5rsneg
  Mac: http://tinyurl.com/nj4m8yt
  Linux: http://tinyurl.com/olb5du2
- Your endpoint will be USERNAME#YOURMACHINENAME

SciNet compute • calcul CANADA

Screenshot 3



Screenshot 4



HPSS

HPSS: longer term storage

- A tape backed hierarchical storage system
- Repository for data that is not actively used.
- Data can be returned to SciNet file system when needed.
- Data movement under control of the user, mostly by scripts.

How to use

- Runs separate from the clusters
- One must either submit a job to push or pull data, using utilities below, or run in a short interactive session

Utilities

- HSI: client with an ftp-like functionality. Can also browse.
- HTAR: creates tar archives directly on HPSS, as well as an index.
- ISH: a TUI utility to perform inventory of your files and tarballs on HPSS.

Interactive session

```
gpc01-$ qsub -q archive  -I
qsub: waiting for job to start
...
hpss-archive01-ib0:~$ pwd
/home/s/scinet/rzon
hpss-archive01-ib0:~$ echo $ARCHIVE
/archive/s/scinet/rzon
hpss-archive01-ib0:~$ hsi
*************************************************************
*     Welcome to HPSS@SciNet - High Perfomance Storage System    *

*                                                               *

*         Contact Information: support@scinet.utoronto.ca       *

*  NOTE: do not transfer SMALL FILES with HSI. Use HTAR instead  *

*               CHECK THE INTEGRITY OF YOUR TARBALLS            *
*************************************************************
[HSI]/archive/s/scinet/rzon->
```

Example script

```
#PBS -l walltime=72:00:00
#PBS -q archive
echo "Creating a htar of finished-job1/ directory tree into HPSS"
trap "echo 'Job script not completed';exit 129" TERM INT
DEST=$ARCHIVE/finished-job1.tar
hsi ls $DEST &> /dev/null
status=$?
if [ $status == 0 ]; then
    echo 'File $DEST already exists. Nothing has been done'
    exit 1
fi
cd $SCRATCH/workarea/
htar -Humask=0137 -cpf $ARCHIVE/finished-job1.tar finished-job1/
status=$?
trap - TERM INT
if [ ! $status == 0 ]; then
   echo 'HTAR returned non-zero code.'
   /scinet/gpc/bin/exit2msg $status
   exit $status
fi
```

Example script

```
#PBS -l walltime=72:00:00
#PBS -q archive
echo "Creating a htar of finished-job1/ directory tree into HPSS"
trap "echo 'Job script not completed';exit 129" TERM INT
DEST=$ARCHIVE/finished-job1.tar
hsi ls $DEST &> /dev/null
status=$?
if [ $status == 0 ]; then
    echo 'File $DEST already exists. Nothing has been done'
    exit 1
fi
cd $SCRATCH/workarea/
htar -Humask=0137 -cpf $ARCHIVE/finished-job1.tar finished-job1/
status=$?
trap - TERM INT
if [ ! $status == 0 ]; then
   echo 'HTAR returned non-zero code.'
   /scinet/gpc/bin/exit2msg $status
   exit $status
fi
```

Data Plan

Data plan

Before starting a computational project, you should have a data plan:

- What is the data you need?
- What will you produce? How much, how many files, organization?
- What can you delete?
  Consider what is faster: recompute or reclaim from tape?
- What of that do you need to keep, and for how long?
- In what form (tape, live, . . .)?
  Don't use many small files. Write in binary. Read in big chunks. Minimize I/O where you can.
- What when your project is done?
- What when you move jobs?

SciNet compute • calcul CANADA

Conclusions

Data is not trivial, but important.

There are many tools and possible workflows.

We can help if you need it.

support@scinet.utoronto.ca