Data Management and Transfer Tools available at SciNet

Ramses van Zon

SciNet HPC Consortium

February 11, 2015



1 / 35

February 11, 2015

Ramses van Zon (SciNet HPC Consortium)Data Management and Transfer Tools availe

Introduction



2 / 35

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

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

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

Outline

- Storage Places
- 2 Data (re)organization
- Moving Data
- Oata plan





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



• \$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.



7 / 35

February 11, 2015

- \$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.



- \$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.



- \$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.



- \$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.



- \$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:

ср х у	#	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 t	tarb	all y	.tar	
tar	caf y.tar.gz a	#	pack+compress	folder	r a	into	tarball	y.tar.gz
zip	y.zip a	#	pack+compress	folder	r 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	y lis	stin	ng of t	arval	conter	ıt
ish	index y.tar.gz	#	crea	ite i	inde	ex file	2		
ish	y.tar.gz.igz	#	ls,	du,	cd	inside	e 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 ware

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

ail 🖹 cc 🗎 s	cinet muot ms r fb ♦ weer 🙆 su 😌 db 🛐 cal m tv m In 🖉 coursera 📿 github m Other E	00
2	globus Manage Data Groups Support rzon r	
	Transfer Files Activity Manage Endpoints Dashboard	
-	Fransfer Files Get Globus Connect Personal Turn yeur computer into an endpoint. Turn yeur computer into an endpoint.	
	Endpoint computecanada#gpc ···· Go S Endpoint enter endpoint name ··· Go	
	Path Go Path Go	
	Please authenticate to access this endpoint	
	When you click CONTINUE you will be refereded to the andpoint login webpage (you will be returned here once you've authenticated). Please select an endpoint above.	
	Continue Cancel	

Screenshot 2

c 🛅 scinet 🛅 uot 🛅 \$ 😭 fb 🔶 weer 📀 su 🔤 db	8 cal 🛅 tv 🛅 In 🕑	coursera 🎧 github	🛅 Oth
🞐 globus		Manage Data Groups S	
	Trans	fer Files Activity Manage Endpo	ints Dashboard
Transfer Files		Get Glo Turn your oo	bus Connect Personal mputer into an endpoint.
Endpoint computecanada#gpc Go		Endpoint enter endpoint name	Go
Path /~/Code/ Go		Path	Go
eelect all none up one folder C refresh list 1. TOOLS 2. PROCRAMS 4. TESTS 9. R EXT Fortran PHOS 9. PYTHON SONET Variest anyaray bashcolpt code	Folder Folder Folder Folder Folder Folder Folder Folder Folder Folder Folder Folder Folder	Please select an endpoint abc	ve.
coa c	Folder Folder Folder Folder Folder		

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



Screenshot 3



Screenshot 4

🔊 globus		Manage Data Groups	
		Transfer Files Activity Manage Endp	oints Dashboard
Transfer Files		Get G Tum your	lobus Connect Personal computer into an endpoint.
Endpoint computecanada#gpc	Go	Endpoint rzon#neptune	Go
Path /~/Code/ Go		Path /~/	Go
select all none t up one folder 🖒 refresh list	=	select all none 눈 up one folder 🖒 refresh lin	st =
1 TOOLS	Folder 🔺	Backup	Folder
2 PROGRAMS	Folder	Canopy	Folder
4 TESTS	Folder	Code	Folder
5 LIBRARIES	Folder =	Desktop	Folder
9_R	Folder	Downloads	Folder
EXT .	Folder	Dropbox	Folder
Fortran	Folder	Enthought	Folder
PHGS	Folder	Notes	Folder
PYTHON	Folder	Coperations	Folder
SCINET	Folder	Personal	Folder
Viar-test	Folder	Pictures	Folder
anyarray	Folder	Projects	Folder
bashscripts	Folder	Publications	Folder
code	Folder	R R	Folder
euda 🛑	Folder	Reference	Folder
hard-spheres	Folder	Teaching	Folder
ideal-gas	Folder	Templates	Folder
intro-gpu	Folder	Tests	Folder
laptop.obj	Folder	Travel	Folder
📫 📫 U	Folder 👻	Mebcam	Folder 🗨

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 *	
	***************************************	u
	[HSI]/archive/s/scinet/rzon->	
Ra	mses van Zon (SciNet HPC Consortium)Data Management and Transfer Tools availe February 11, 2015 31 / 35	j

Example script

```
#PBS -1 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 -1 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?

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

