

UF Research Computing: Overview and Running STATA



Matt Gitzendanner
maqitz@ufl.edu

9/24/14


 Home of High-Performance Computing and **HiPerGator**

 Information Technology 

Research Computing







- ▶ Mission
 - Improve opportunities for research and scholarship
 - Improve competitiveness in securing external funding
 - Provide high-performance computing resources **and support** to UF researchers



 Information Technology 


Research Computing



- ▶ Funding
 - Faculty
 - Matching grant program!
- ▶ Any UF Faculty can use
 - Up to 8 cores
 - Investors gain priority and access to additional resources
- ▶ Comprehensive management
 - Hardware maintenance and 24x7 monitoring
 - Relieve researchers of the majority of systems administration tasks



 Information Technology 

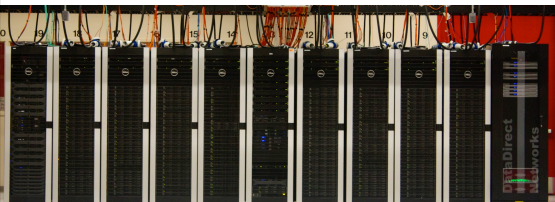



 Information Technology 

HiPerGator

The University of Florida Supercomputer for Research


- 16,384 cores—*total of about 21,000 cores today*
- Infiniband interconnect
- >3PB fast, high-availability, storage
- **GPGPUs**
- Large memory nodes (**512GB to 1TB of RAM**)





 Information Technology

Research Computing

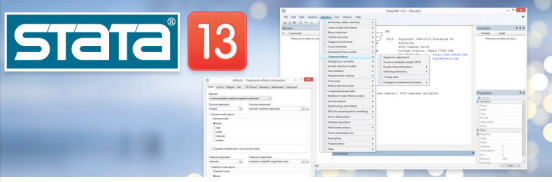
- ▶ User Accounts
 - Qualifications:
 - Current UF faculty sponsor
 - Non-investors can use 8 cores at a time
 - Investors gain priority access



www.rc.ufl.edu

 Information Technology 

Research Computing

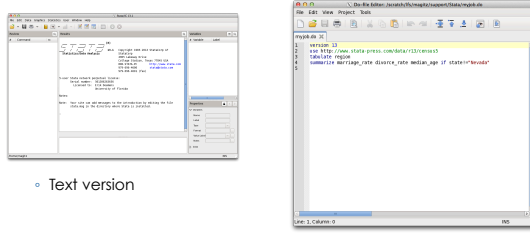


Running STATA

UF Information Technology www.it.ufl.edu

STATA 13

- Interactive
 - GUI version
- Submit a STATA do file
- Text version



UF Information Technology www.it.ufl.edu

STATA 13

- GUI Versions available
 - xstata
 - xstata-mp
 - xstata-se
 - xstata-sm
- Non-GUI versions
 - stata
 - stata-mp
 - stata-se
 - stata-sm


X for X11

-mp: Multiprocessor (8-core license)
 -se: Special Edition
 -sm: ??

UF Information Technology www.it.ufl.edu

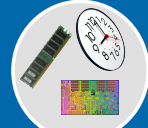
Cluster Basics

User interaction




Login server (Head node)

Scheduler



Tell the scheduler what you want to do

Compute resources



Your job runs on the cluster


UF Information Technology www.it.ufl.edu

Cluster Basics

Interactive nodes

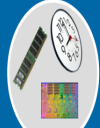
- dev1
- dev2
- gui
- gui1

User interaction




Login server (Head node)

Scheduler



Tell the scheduler what you want to do

Compute resources



Your job runs on the cluster

UF Information Technology www.it.ufl.edu


Cluster login—for Text or do file

Hostname: **gator.rc.ufl.edu**


Mac/Linux in Terminal:
`ssh user@gator.rc.ufl.edu`

Windows PuTTY:
 Hostname: gator.rc.ufl.edu

Windows: PuTTY
 Mac/Linux: Terminal



User interaction






Login server (Head node)

UF Information Technology www.it.ufl.edu


Cluster login – GUI Nodes

For GUI nodes:
`ssh -Y <user>@gui.rc.ufl.edu`
 or
`ssh -Y <user>@gui1.rc.ufl.edu`

Windows: PUTTY 
 Mac/Linux: Terminal 

Windows: Xming or MobaXterm 
 Mac: XQuartz

User interaction



Login server (Head node)

UF Information Technology www.it.ufl.edu

Development servers

- Do not run applications on the login servers
 - Account will be suspended

Do not run interactive jobs on the login nodes.

UF HPC Center Account Policies can be found here:
<http://www.hpc.ufl.edu/about/policies/account>
`[magitz@gator1 ~]$`

- Use the development servers for testing and interactive use:
 - `ssh dev1` or `ssh dev2`


UF Information Technology www.it.ufl.edu

Cluster Storage

gator.rc.ufl.edu

gator1	gator3
gator2	gator4

- /home/\$USER: 20GB limit**
 - Do not use for input or output for jobs
- /scratch/lfs/\$USER: 2TB/group**
 - Primary storage for all HiPerGator jobs



UF Information Technology www.it.ufl.edu

Helpful tips

- Don't put spaces in file/folder names
- Use a text editor to save Unix line breaks
 - Or run: `dos2unix filename`

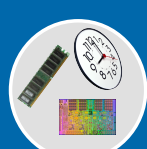
```
$ dos2unix sample_code.do
dos2unix: converting file sample_code.do to UNIX format ...
$
```

UF Information Technology www.it.ufl.edu

Scheduling a job

- Need to tell scheduler what you want to do
 - How many CPUs** you want and how you want them grouped
 - How much RAM** your job will use
 - How long** your job will run
 - The commands that will be run

Scheduler



Tell the scheduler what you want to do

UF Information Technology www.it.ufl.edu

Submit script for STATA do file

```
#!/bin/bash
#PBS -N My_Job_Name
#PBS -M Joe_Shmoie@ufl.edu
#PBS -m abe
#PBS -o My_Job.log
#PBS -e My_Job.err
#PBS -l nodes=1:ppn=1
#PBS -l pmem=900mb
#PBS -l walltime=00:05:00

cd $PBS_O_WORKDIR

module load stata
stata my_job.do
```

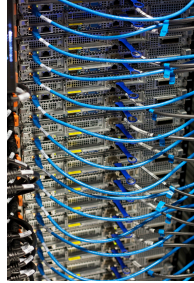
Mostly optional info

Resource request

UF Information Technology www.it.ufl.edu

Nodes and processors

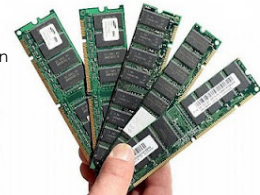
- For STATA nodes=1 always
- For stata, stata-se:
#PBS -l nodes=1:ppn=1
- For stata-mp
#PBS -l nodes=1:ppn=8
We are licensed for up to 8-cores with stata-mp



RAM

#PBS -l pmem=900mb

- **Per-processor** memory request
- Lots to consider, but do your best at estimating memory needed for job
- Over about 4GB of RAM, "costs" toward CPU allocation



Wasted RAM leads to idle CPUs and low job throughput

End-of-job emails: #PBS -M Joe_Shmoe@ufl.edu #PBS -m abe

```
PBS Job Id: 358634.moab.ufhpc
Job Name: NR.25.nex
Exec host: c7a-s1/60
Execution terminated
Exit_status=0
resources_used.cput=07:16:09
resources_used.mem=251348kb
resources_used.vmem=318916kb
resources_used.walltime=07:16:52
```

End-of-job emails: #PBS -M Joe_Shmoe@ufl.edu #PBS -m abe

```
PBS Job Id: 5671727.moab.ufhpc
Job Name: muscle
Exec host: s5a-s23/0
job deleted
Job deleted at request of root@moab.ufhpc
job 5671727 exceeded MEM usage hard limit
(19851 > 18022)
```

Example of a job that was deleted because it exceeded the memory request

Walltime

#PBS -l walltime=00:50:00

- Fairly straight forward
- As with all resource requests, accuracy helps ensure **your** jobs and all other jobs will run sooner



	Maximum	Short	Long
Investor	31 days	<12 hrs	7 days
Other	7 days	<12 hrs	3 days

Research Computing

- Job Management
 - qsub <file_name>: job submission
 - qstat -u <user>: check queue status
 - showq -r -u <user>: shows job efficiency
 - qdel <JOB_ID>: job deletion
 - checkjob -v <job number> (shows PE value)
 - pbs_info -f my_job.pbs (get job PE and group resources before submitting a job)

UF Research Computing

▸ Help and Support

- <https://support.rc.ufl.edu>
 - For any kind of question or help requests
- <http://wiki.rc.ufl.edu>
 - Documents on hardware and software resources
 - Various user guides
 - Many sample submission scripts
- <http://rc.ufl.edu>
 - Frequently Asked Questions
 - Account set up and maintenance

