



## UF Research Computing: An Introduction and Getting Started



Matt Gitzendanner  
[mgqitz@ufl.edu](mailto:mgqitz@ufl.edu)

9/3/14

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

UF Information Technology
www.it.ufl.edu

## Research Computing




**UF** Research Computing  
Information Technology  
Home of High-Performance Computing and **HiPerGator**

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

UF Information Technology
www.it.ufl.edu

## 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



UF Information Technology
www.it.ufl.edu


## Matching Program




# \$1.4 million

UF Information Technology
www.it.ufl.edu

UNIVERSITY OF FLORIDA | High-Performance Computing

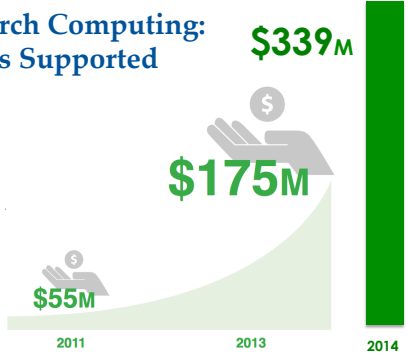


# HiPerGator

The University of Florida Supercomputer for Research

UF Information Technology
www.it.ufl.edu

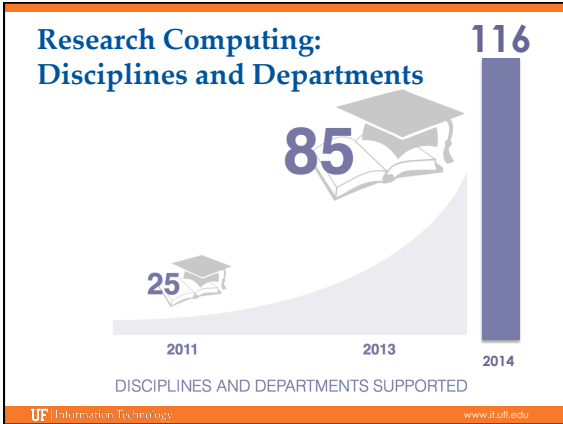
## Research Computing: Grants Supported



2011: \$55M  
2013: \$175M  
2014: \$339M

GRANTS SUPPORTED

UF Information Technology
www.it.ufl.edu



## 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)

## Research Computing

Where do you start?

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)


## 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](http://www.rc.ufl.edu)

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

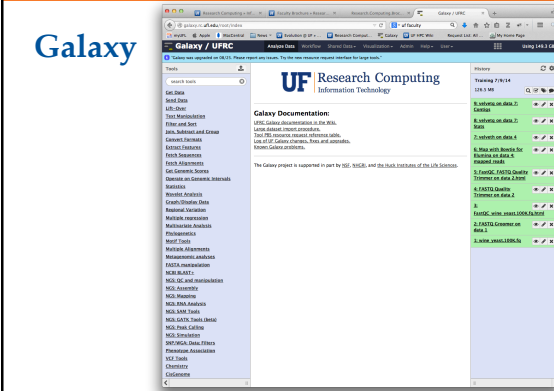
## What can you run?



- ▶ Linux-based
- ▶ Generally command line driven applications
- ▶ Galaxy 
- ▶ Graphical apps can be setup
  - SAS 
  - MATLAB 

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Galaxy



UF Research Computing Information Technology

Galaxy Documentation: UFRC Galois documentation is the Web-based Galaxy interface. It is used for Galaxy workflow management, and is the main interface for Galaxy users. It is the main interface for Galaxy users. It is the main interface for Galaxy users.

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Galaxy

Map with Bowtie for Illumina

Will you select a reference genome from your history or use a built-in index?

Built-ins were indexed using default options

Select a reference genome:

If your genome of interest is not listed - contact Galaxy team

Is this library mate-paired?

FASTQ file:

Must have ASCII encoded quality scores

Bowtie settings to use:

For most mapping needs use Commonly used settings. If you want full control use Full parameter list

Suppress the header in the output SAM file:  
 Bowtie produces SAM with several lines of header information by default

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Wiki.rc.ufl.edu



UF Research Computing Wiki

Welcome to the University of Florida Research Computing Wiki. The information here augments what is found on our web site with information we think is best provided in a less formal and official context. It is used for information that changes rapidly and might become quickly dated or incorrect on the web site. While there is good and helpful information here, you will find text that has not been proofed to the same standards that we try to maintain on our web site. We hope you find it helpful.

Getting Started

- Getting Started
- Mailing Lists
- Changing your Password
- Moab Command Summary
- Training

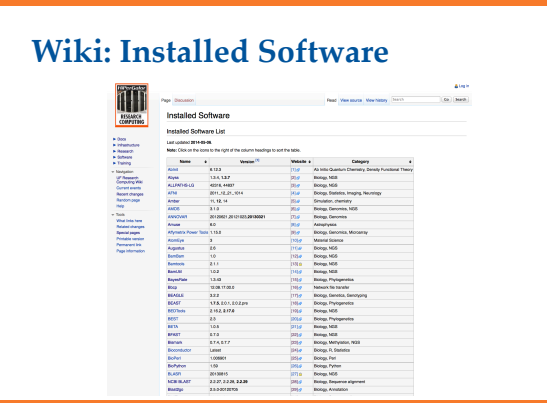
Software and Libraries

- Installed Software
- Environment Modules System
- Installing Perl Modules
- GPUs and CUDA

Connecting and Data Transfer

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Wiki: Installed Software

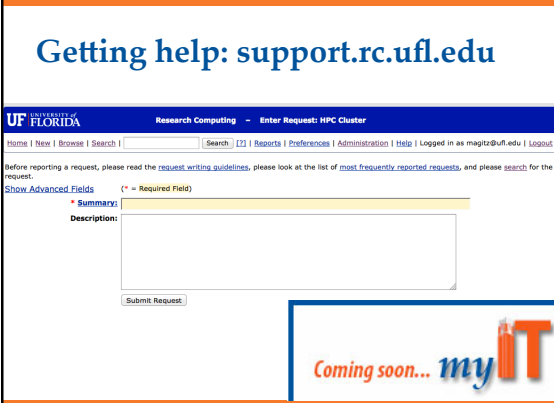


Installed Software

Name	Version	Category
AbiDoc	1.0.1	Image
AbiGen	1.0.1	Image
AbiView	1.0.1	Image
AbiView2	1.0.1	Image
AbiView3	1.0.1	Image
AbiView4	1.0.1	Image
AbiView5	1.0.1	Image
AbiView6	1.0.1	Image
AbiView7	1.0.1	Image
AbiView8	1.0.1	Image
AbiView9	1.0.1	Image
AbiView10	1.0.1	Image
AbiView11	1.0.1	Image
AbiView12	1.0.1	Image
AbiView13	1.0.1	Image
AbiView14	1.0.1	Image
AbiView15	1.0.1	Image
AbiView16	1.0.1	Image
AbiView17	1.0.1	Image
AbiView18	1.0.1	Image
AbiView19	1.0.1	Image
AbiView20	1.0.1	Image
AbiView21	1.0.1	Image
AbiView22	1.0.1	Image
AbiView23	1.0.1	Image
AbiView24	1.0.1	Image
AbiView25	1.0.1	Image
AbiView26	1.0.1	Image
AbiView27	1.0.1	Image
AbiView28	1.0.1	Image
AbiView29	1.0.1	Image
AbiView30	1.0.1	Image
AbiView31	1.0.1	Image
AbiView32	1.0.1	Image
AbiView33	1.0.1	Image
AbiView34	1.0.1	Image
AbiView35	1.0.1	Image
AbiView36	1.0.1	Image
AbiView37	1.0.1	Image
AbiView38	1.0.1	Image
AbiView39	1.0.1	Image
AbiView40	1.0.1	Image
AbiView41	1.0.1	Image
AbiView42	1.0.1	Image
AbiView43	1.0.1	Image
AbiView44	1.0.1	Image
AbiView45	1.0.1	Image
AbiView46	1.0.1	Image
AbiView47	1.0.1	Image
AbiView48	1.0.1	Image
AbiView49	1.0.1	Image
AbiView50	1.0.1	Image

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Getting help: support.rc.ufl.edu



UF UNIVERSITY OF FLORIDA Research Computing - Enter Request: HPC Cluster

Before reporting a request, please read the request writing guidelines, please look at the list of most frequently reported requests, and please search for the request.

Show Advanced Fields (\* = Required Field)


Summary:

Description:

Coming soon... myIT

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

UNIVERSITY OF FLORIDA | High-Performance Computing



# HiPerGator

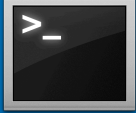
The University of Florida Supercomputer for Research

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Cluster Basics

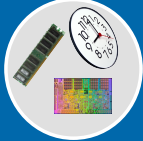
User interaction

**Galaxy**




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](http://www.it.ufl.edu)

## Cluster login

ssh → hipergator.rc.ufl.edu

gator1

gator3


gator2

gator4


/home/\$USER

ssh <user>@gator.rc.ufl.edu


Windows: PuTTY



Mac/Linux: Terminal



User interaction



Login server (Head node)

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Cluster login

ssh → hipergator.rc.ufl.edu


gator1 gator3

gator2 gator4

/home/\$USER

ssh <user>@gator.rc.ufl.edu

User interaction



Login server (Head node)

```

FLMNH-SOL-MAC1:~ gitzs ssh magitz@gator.rc.ufl.edu
magitz@gator.rc.ufl.edu's password:
Last login: Mon May 19 09:37:33 2014 from 10.243.21.61

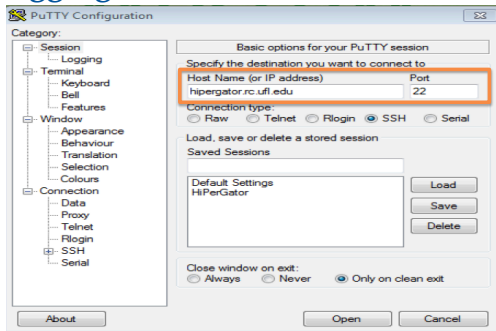
Welcome to the UF HPC Center.

Do not run interactive jobs on the login nodes. If you need to
run an interactive job, there are interactive/test nodes for that.

s http://wiki.hpc.ufl.edu/doc/Test_Nodes
W
M http://www.hpc.ufl.edu/about/policies/account
[magitz@gator4 ~]$
    
```

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Logging in PuTTY



UF Information Technology [www.it.ufl.edu](http://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


User interaction



Login

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

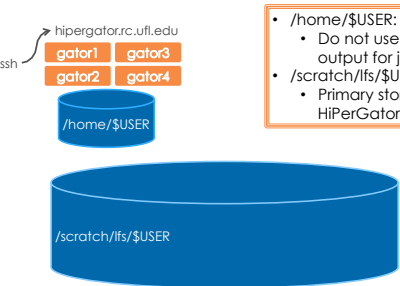
## Linux Command Line



- ▶ Lots of online resources
  - Google: Linux cheat sheet
  - ▶ Training sessions
  - Sept. 4: The Linux/Unix Command Line - An Introduction
  - ▶ User manuals for applications

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Cluster Storage



- /home/\$USER: 20GB limit
  - Do not use for input or output for jobs
- /scratch/lfs/\$USER: 1TB limit
  - Primary storage for all HiPerGator jobs

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Research Computing

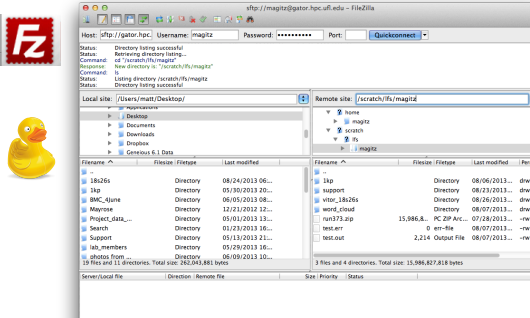


- ▶ Storage
  - **Home:** /home/\$USER
    - For code compilation and user file management only
    - **Do not use for job input/output!**
    - Include `cd $PBS_O_WORKDIR` or similar in scripts
  - **Scratch space:** Lustre File System
    - /scratch/lfs/\$USER

Other storage options available for purchase

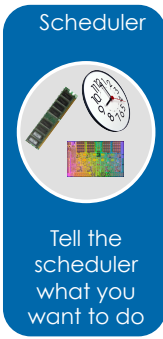
UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## SFTP-e.g.: FileZilla or CyberDuck



UF Information Technology [www.it.ufl.edu](http://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

Tell the scheduler what you want to do

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Research Computing

- ▶ Ordinary Shell Script

```
#!/bin/bash

date
module load test_app
test_app -i file.txt
```

Read the manual for your application

Commands typed on the command line can be put in a script

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

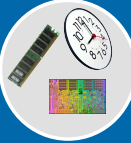
## Research Computing

### Submission Script

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


cd $PBS_O_WORKDIR
date
module load test_app
test_app -i file.txt
```

Scheduler



Tell the scheduler what you want to do

Compute resources




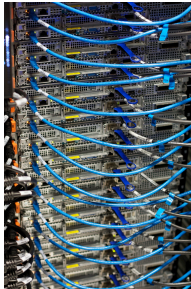
Your job runs on the cluster

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Nodes and processors

```
#PBS -l nodes=1:ppn=4
#PBS -l nodes=2:ppn=32
```






UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## 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

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## End-of-job emails:

```
#PBS -M Joe_Shmoefufl.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
```

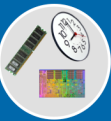
UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## 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

Scheduler



Tell the scheduler what you want to do

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

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## 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 Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

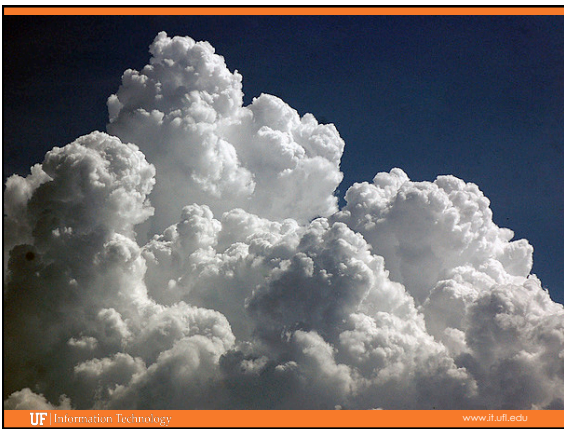
## Research Computing

- ▶ Job Scheduling and Usage
  - Job scheduler selects jobs based on priority
    - Priority is determined by several components
    - Investors have higher priority
    - Non-investor jobs limited to 8 processor equivalents (PEs)
    - RAM: requests beyond a few GB/core start counting toward the total PE value of a job
  - Test nodes (dev1 and dev2) available for interactive use, testing and short jobs
  - Limit number of jobs to 2-3,000 at a time.
    - Combine multiple short jobs into a single job.

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## UF RISING to National Preeminence

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)



## 2014 Projects

- ▶ GatorCloud
  - OneDrive@UF
    - 1TB of document and file storage
    - Access to MS Office Online web applications (Word, Excel, PowerPoint, and OneNote)
    - Approved storage of UF restricted data, excluding Protected Health Information (PHI)
  - GatorWOS/Box
  - GatorVault
  - UF Research Apps
  - GatorCloud Virtual Systems
  - GatorCloud Portal

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## 2014 Projects

- ▶ UF Research Apps

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## 2014 Projects

- ▶ Virtualization
  - GatorCloud—Dr. Andy Li

UF Information Technology [www.it.ufl.edu](http://www.it.ufl.edu)

## Training Schedule

- ✓ Aug 28: Intro to UFHPC, getting started
- ▶ Sep 4: The Linux/Unix Shell - An Introduction
- ▶ Sep 11: HiPerGator: Running Jobs, Submission Scripts, Modules
- ▶ Sep 18: MATLAB at Research Computing: A Hands-on Tutorial to Running Your MATLAB Code at Research Computing

More training sessions  
will be added

## 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

