

UF Research Computing: An Introduction and Getting Started

Matt Gitzendanner
mqgitz@ufl.edu

1/23/14

UF Research Computing
Information Technology
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

Consolidating Resources to Improve Efficiency and Capacity




The Research Computing Matching Program pooled \$642k. Thereby creating synergies and improving research infrastructure.

2011

UF | Information Technology | www.it.ufl.edu

UNIVERSITY OF FLORIDA | High-Performance Computing

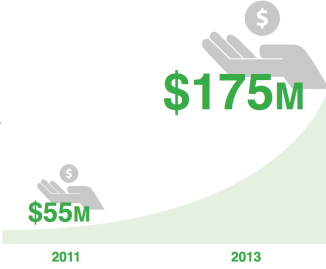


HiPerGator

The University of Florida Supercomputer for Research

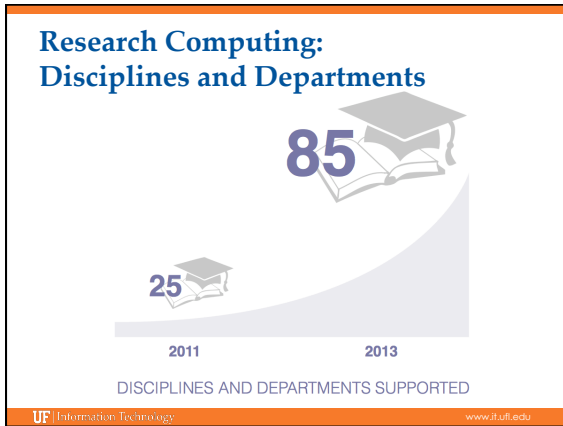
UF | Information Technology | www.it.ufl.edu

Research Computing: Grants Supported



GRANTS SUPPORTED

UF | Information Technology | www.it.ufl.edu



Internet2 Network

UF was 1st in the nation

- Internet2 Innovation Platform
 - 100 Gbps connectivity
 - Campus Research Network now 200 Gbps

A photograph showing rows of server racks in a data center. The racks are illuminated with blue and yellow lights, creating a futuristic atmosphere.

UF Information Technology www.it.ufl.edu

UF Data Center

- UF Data Center on Eastside Campus
 - 10,000 sq.ft and 1.75 MW total
 - 5,000 sq. ft. space for Research Computing

A photograph of the exterior of a brick building with two gabled roofs and small windows. The sky is blue with some clouds.

UF Information Technology www.it.ufl.edu

HiPerGator

The University of Florida Supercomputer for Research

- 16,384 cores—total of about 19,000 cores today
- Infiniband interconnect
- >3PB fast, high-availability, storage
- GPGPUs— 90+, Kepler-class
- Large memory (512GB to 1TB of RAM) nodes

A photograph of the interior of a server room, showing rows of black server racks with various components and cables.

UF Information Technology www.it.ufl.edu

Research Computing

A square icon representing a terminal window with a black background and white text showing a prompt character and a cursor.

Where do you start?

UF Information Technology 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

A screenshot of the UF Research Computing website, showing a 'Request Account' form with various input fields and a 'Request' button.

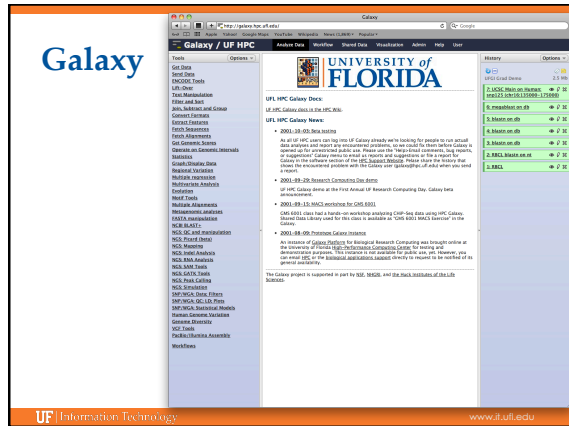
www.rc.ufl.edu

UF Information Technology www.it.ufl.edu

What can you run?

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

UF Information Technology www.it.ufl.edu



Galaxy / UF HPC

UF Information Technology www.it.ufl.edu

Galaxy: Data intensive biology for everyone

- ▶ Accessible, reproducible, transparent computational biology
- ▶ galaxy.hpc.ufl.edu
 - Local instance of Galaxy
 - Faster access to storage, easier upload
 - Local compute resources
 - Local control

UF Information Technology 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



Galaxy

UF Information Technology www.it.ufl.edu

What can you run?

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

UF Information Technology www.it.ufl.edu

Logging in

UF Information Technology www.it.ufl.edu

Development servers

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

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.

http://wiki.hpc.ufl.edu/doc/Test_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 server (Head node)

UF Information Technology www.it.ufl.edu

Linux Command Line

- Lots of online resources
 - Google: Linux cheat sheet
 - Training sessions
 - Jan 23: The Linux/Unix Command Line - An Introduction
 - User manuals for applications

UF Information Technology www.it.ufl.edu

Cluster Storage

- /home/\$USER: 10GB 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

SFTP-e.g.: FileZilla or CyberDuck

UF Information Technology www.it.ufl.edu

Research Computing

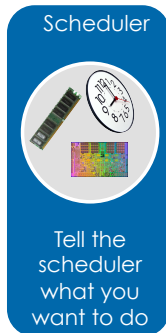
- Storage
 - Home: /home/\$USER
 - For code compilation and user file management only
 - Do not use for job input/output!
 - Include cd \$PBS_O_WORKDIR or similar in scripts
 - Scratch space: Lustre File System
 - /scratch/lfs/\$USER

Other storage options available for purchase

UF Information Technology www.it.ufl.edu

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



Research Computing

- ▶ Ordinary Shell Script

```
#!/bin/bash
date
module load test_app
test_app -i file.txt
```

Read the manual
for your application

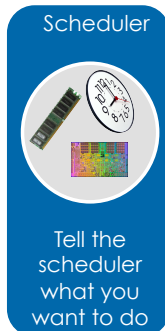
Commands typed
on the command
line can be put in a
script

Research Computing

- ▶ Submission Script

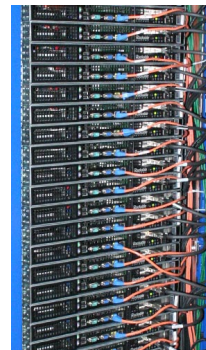
```
#!/bin/bash
#PBS -N My_Job_Name
#PBS -M Joe_Shmoefufl.edu
#PBS -m abe
#PBS -o My_Job_Name.log
#PBS -e My_job_Name.err
#PBS -l nodes=1:ppn=1
#PBS -l walltime=00:05:00
#PBS -l pmem=900mb

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



Nodes and processors

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



RAM

```
#PBS -l pmem=900mb
```

- ▶ **Per-processor** RAM request

- ▶ Lots to consider, but do your best at estimating RAM 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_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
```

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 |

Scheduler



Tell the scheduler what you want to do

UF Information Technology 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

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
 - Connect from login node:


```
[magitz@gator1 ~]$ssh dev1
```

UF Information Technology www.it.ufl.edu

2014 Projects

- Cloud Storage—"GatorBox"
 - Dropbox-like sharing and collaboration
 - Replication across state






SSERCA.org

UF Information Technology www.it.ufl.edu

2014 Projects

- UF Apps for Research








Campus Network, Network Services
Contact: 391.2247


UF Information Technology www.it.ufl.edu

2014 Projects


- Virtualization
 - GatorCloud—Dr. Andy Li

- Data Transfer—Globus
 -



UF Information Technology www.it.ufl.edu



UF RISING
to
National Preeminence

UF Information Technology www.it.ufl.edu

Training Schedule

- ✓ Jan 9: Intro to UFHPC, getting started
- ▶ Jan 23: The Linux/Unix Shell - An Introduction
- ▶ Jan 30: HiPerGator: Running Jobs, Submission Scripts, Modules
- ▶ Feb 13: Galaxy Overview, The Basics
- ▶ Feb 20: NGS: Reference Based Mapping & de Novo Assembly
- ▶ Multiprocessor Applications, GPU usage, Application profiling and Debugging, Storage, Data Transfer

What topics would be most helpful for you??
Email: [magitz@ufl.edu](mailto:magitiz@ufl.edu)

UF Information Technology

www.it.ufl.edu

UF Research Computing

- ▶ Help and Support
 - <https://support.hpc.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



UF Information Technology

www.it.ufl.edu