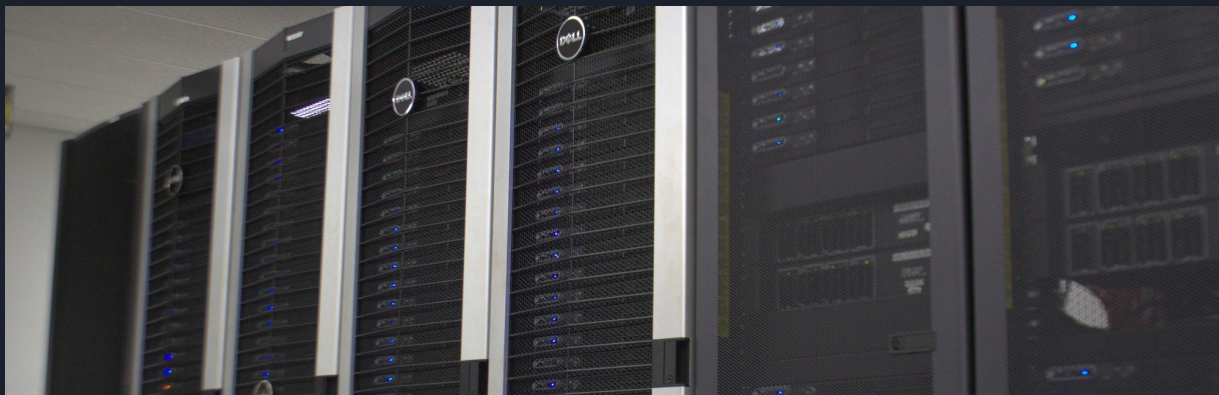


# High-Performance Computing at FSU with the Research Computing Center

Alex Townsend and Marcelina Nagales

# What is the FSU Research Computing Center (RCC)

- A unit of the FSU Information Technology Services division
- Originally a division of Scientific Computing back in 2007
- The team that hosts and administers the main cluster at FSU, the High Performance Computing system (HPC)



# Speaker Introductions



Marcelina Nagales

Software/Applications Team

Joined in Fall 2021

Also a Masters student in Scientific  
Computing under Dr. Quaife



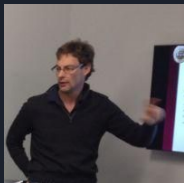
Alex Townsend

Software/Applications Team

Joined in early 2020

Also a PhD student in Scientific  
Computing under Drs. Beerli and  
Meyer-Baese

**Director**



Paul van  
der Mark

**Operations  
Manager**



Alex  
Birkovsky

# The RCC Team

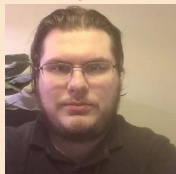
**Support/Applications Team**



Marcelina  
Nagales



Prasad  
Maddumage

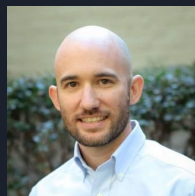


Alex  
Townsend

Digital  
Humanities  
Specialist

We're Hiring!

**Support  
Coordinator**



Casey  
McLaughlin

**Systems/Infrastructure Team**



Cameron  
Berkley



Brian  
Gentry



Mitch  
Gans



Terry  
Ward

# HPC Infrastructure

**808**  
Nodes

Beowulf Cluster  
Centos 8  
AlmaLinux (Maybe)

**32**  
GPUs

NVIDIA GeForce  
GTX1080Ti GPUs  
Future: A30 GPUs

**17,144**  
Processors

AMD, Intel  
Massively Scalable  
Massively Parallel

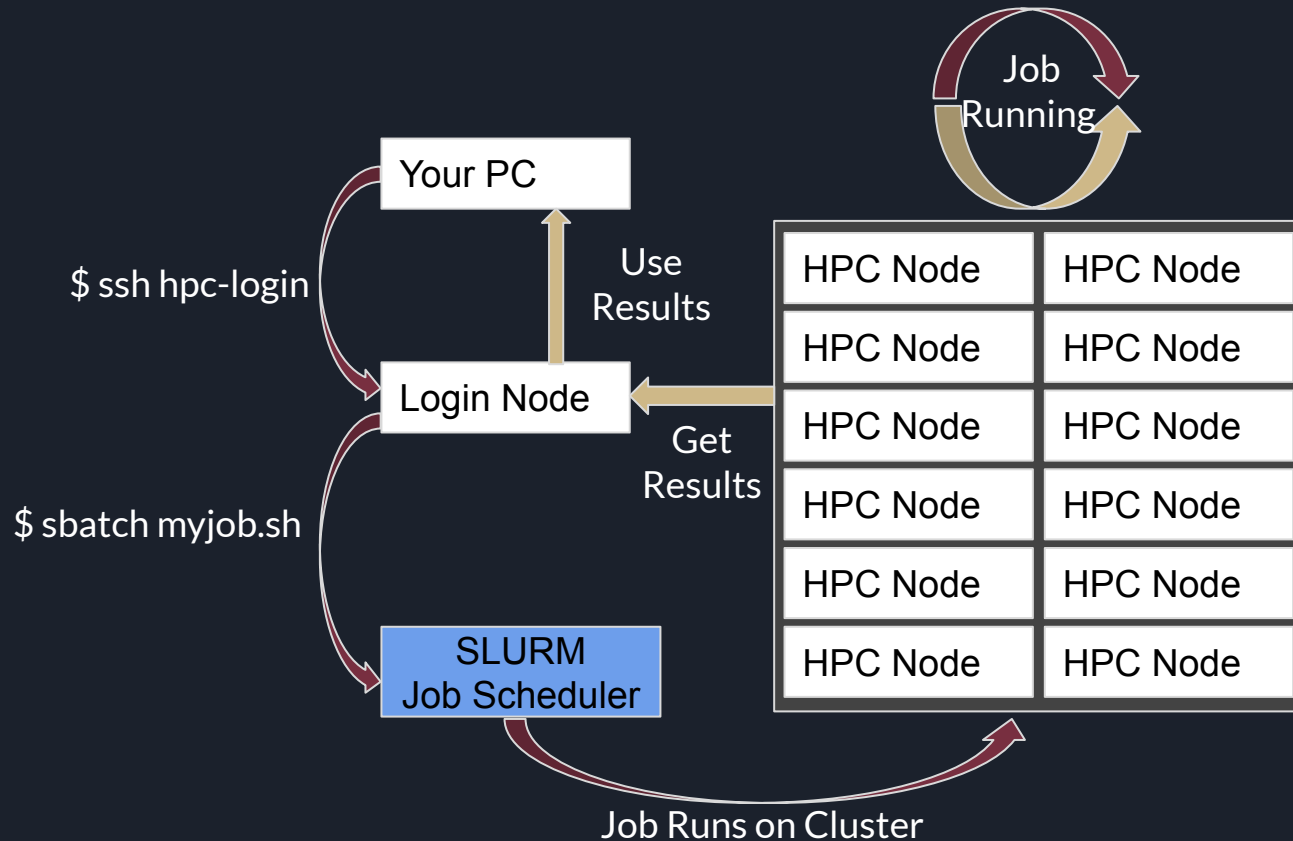
**414,995**  
GigaFLOPS

Very Fast  
Maximum Speeds

**2 Million**  
Jobs Per  
Year (Avg)

Highly Available  
Widely Used

# General HPC Workflow



# How HPC Works

## Job Scheduler: SLURM

- Directs jobs so they can be processed quickly and effectively
- Allocates resources to jobs so the cluster does not freeze up

## Queues/Partitions

- Like “lines” you might wait in when at a grocery store
- Each queue/partition has access to specific HPC nodes
- They process jobs that are submitted to the cluster

# Infrastructure

## Physical Location

- Cluster held in the Sliger Data Center in Innovation Park. Tours available!

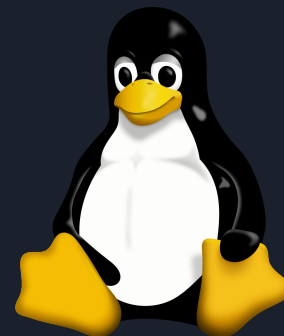
## Standard RCC accounts

- 1 Home directory per user
- 150GB of Parallel Storage
- Temporary Scratch Space (Space Limit is variable)
- Access to the General Access and Backfill Queues
- Accessible via Open On Demand and CLI



## Additional Features for Paid RCC Accounts

- Access to highly scalable Parallel and Archival storage
- Priority access to dedicated computing resources in your own queue
- Custom and specialized hardware and infrastructure (as available)





# Support Services

## Basic Support

- Software installation, configuration, and maintenance
- Technical support for installed applications and software
- Workflow process support and improvement assistance
- Assistance with HPC commands and utilities

## Additional Paid Support ([more Information](#))

- Software development consulting
- In-depth HPC and software focused consulting for research projects

Email us at [support@rcc.fsu.edu](mailto:support@rcc.fsu.edu) for support

# Cybersecurity Initiatives

## System

- Systems are locked behind the FSU VPN
- Systems are secured by a Firewall
- Secure access through VPN
  - requires a username, password and DUO Authentication

## Network

- FSU network traffic is monitored by a central IDS/IDPS
- Networks have multiple layers of access filters
- Physical Access to the servers is restricted by keycard
- Restrictive Default Home and Research Directory Access Permissions

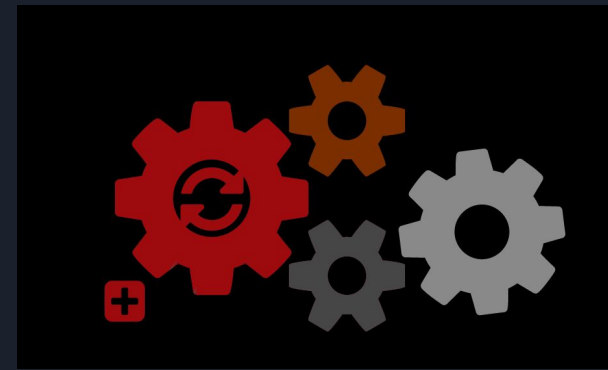
# Maintenance Cycles

## Annual Maintenance

- Requires ~1 week of downtime
- Hardware repair
- Upgrades and maintenance to data center, network, and low-level security
- Vital for maintaining the stability, reliability and security of our systems.

## Annual Software Upgrade

- Requires ~1 week of reduced capacity but not full downtime
- Upgrades to OS, default software, Job Scheduler
- Patches and recompilation
- Vital for maintaining the security and modernity of our systems.



By Muhammad Rafizeldi (MRafizeldi) - Own work, CC BY-SA 3.0,  
<https://commons.wikimedia.org/w/index.php?curid=29381072>

# Support Capabilities

- Regular Office Hours 8AM - 5PM Monday through Friday
  - We work a hybrid schedule but are available by Zoom or email
- Applications Specialists dedicated to software, user and research support
- Systems Specialists dedicated to maintaining and supporting our hardware and infrastructure.
- A strong leadership team with combined over 40 years of deep technical expertise

Email us at [support@rcc.fsu.edu](mailto:support@rcc.fsu.edu) for any questions or concerns

# Primary HPC Use Cases

## Computations

- Highly repetitive, computationally intensive and data-intensive tasks:
  - Parallelized across multiple processors
  - Optimized for distributed computing (i.e. OpenMPI)
  - Requires specialized hardware that can be difficult to acquire or support long-term (i.e. GPUs, DSPs, Vector Co-Processors, etc)
  - Long runtime with minimal chance of interruption
  - Requires a lot of RAM

## Storage

- RCC has several Petabytes worth of storage available across our Parallel and Archival Storage Systems
- Available for long-term storage of data and code

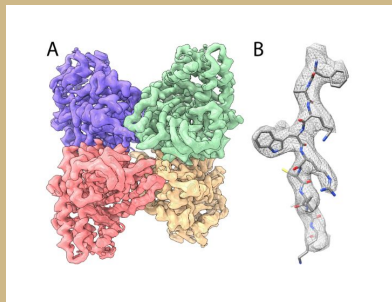
Email us at [support@rcc.fsu.edu](mailto:support@rcc.fsu.edu) for a quote

# Examples of HPC-Enabled Research

## Molecular Biophysics

3D Image Processing of  
Molecular Structures

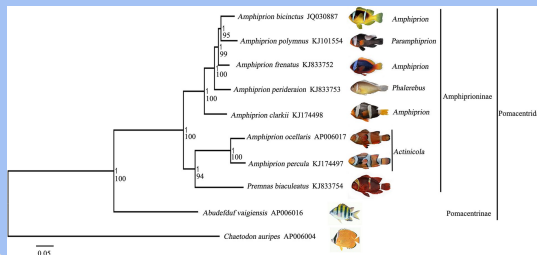
*Dr. Scott Stagg & CryoEM  
Group*



## Population Genetics

Scalable Inference of  
Population Phylogenetic  
Parameters

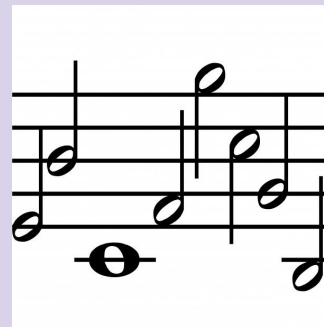
*Dr. Peter Beerli*



## Music

A Statistical Look at Harmony  
in Music

*Daniel Tompkins*



# Examples of HPC-Enabled Research

## EOAS

Hurricane Simulations Help  
Facilitate Forecasts and  
Enhance Preparations

*Dr. Allison Wing*



## Business

Managing Storm Risk at the  
College of Business

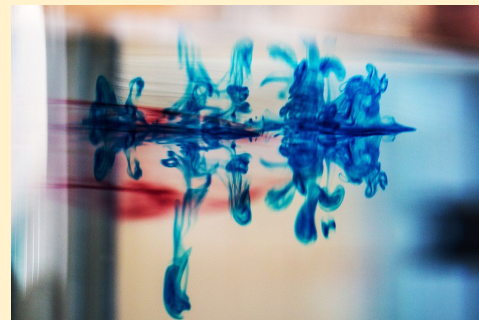
*Sue Ellen Smith*



## Engineering

Simulating and Controlling  
Fluid Flow at AME and RCC

*Kunihiko Taira*



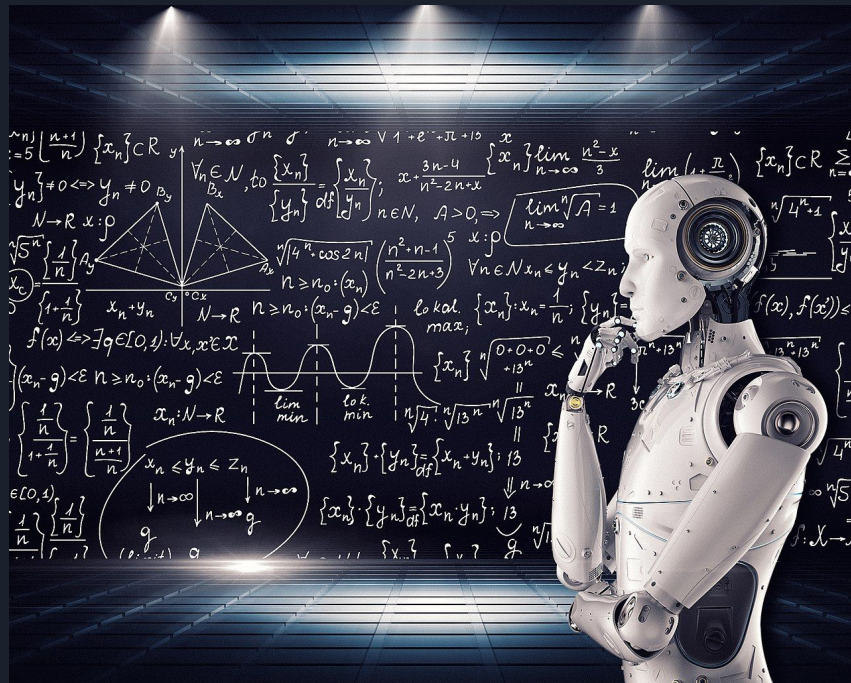
# Data Science and Machine Learning

## Infrastructure

- GPU Nodes
- UROP Projects
- In-House Expertise

## Popular Packages

- TensorFlow, PyTorch, R, Anaconda, CUDA, cuDNN, OpenCV
- Apache Spark

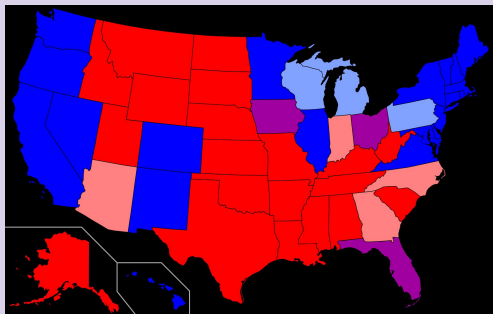




# Data Science and Machine Learning Enabled by HPC

## Sentiment Analysis of the 2016 Presidential Election

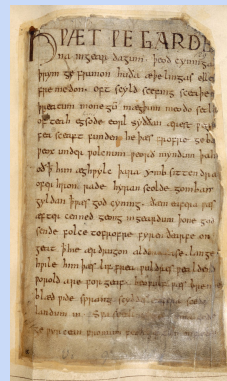
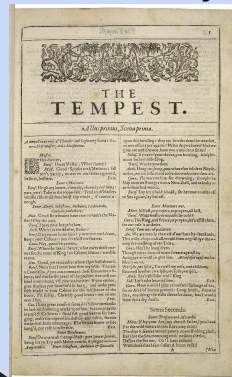
*Dr. Prasad Maddumage*



By Angr - self-made; base map is Image:Blank US Map.svg, CC BY-SA 3.0,  
<https://commons.wikimedia.org/w/index.php?curid=3532565>

## Machine Learning and Shakespeare Works

*Carolyn Linehan*



By William Shakespeare, Isaac Jaggard and Edward Blount  
(printers) - Folger Shakespeare Library Digital Image Collection  
<http://luna.folger.edu/luna/servlet/s/e5fhy8>, CC BY-SA 4.0,  
<https://commons.wikimedia.org/w/index.php?curid=40907202>

# Register For an Account

## Faculty Requirements

- FSU ID and password

## Student Requirements

- FSU ID and password
- Faculty Sponsor

## Guest (non-FSU)

- Guest FSU ID
- Faculty Sponsor



The screenshot shows the FSU Research Computing Center (RCC) login and registration page. The header includes the Florida State University logo and the text "FSU RESEARCH COMPUTING CENTER" and "INFORMATION TECHNOLOGY SERVICES". The page has a navigation bar with "HOME" and "MANAGE" links. The main content area is divided into two columns. The left column has a "Manage" section with links for "Account", "Groups", and "HPC Partitions", and a feedback box. The right column has an "FSUID Login" section with a text input for the FSUID and a password input, and a "Don't have an FSUID?" section with a "Signup as Guest" button. A "Login" button is at the bottom of the login section.

FLORIDA STATE UNIVERSITY

FSU RESEARCH COMPUTING CENTER  
INFORMATION TECHNOLOGY SERVICES

HOME MANAGE

Manage

Account

Groups

HPC Partitions

If you notice any issues or have suggestions, please let us know.

FSUID Login

You can sign-up or login using your FSUID.

FSUID

FSUID Password

(this is the same account you use to login to Canvas or my.fsu.edu)

Login Reset system account password.

Don't have an FSUID?

You need to have an FSUID in order to sign-up for a RCC account. Fortunately, getting a guest FSUID requires only a few steps.

Signup as Guest

[www.rcc.fsu.edu/manage](http://www.rcc.fsu.edu/manage)

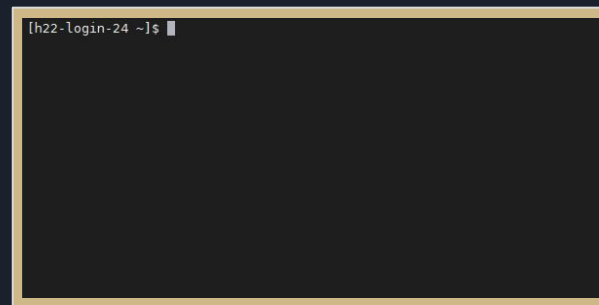
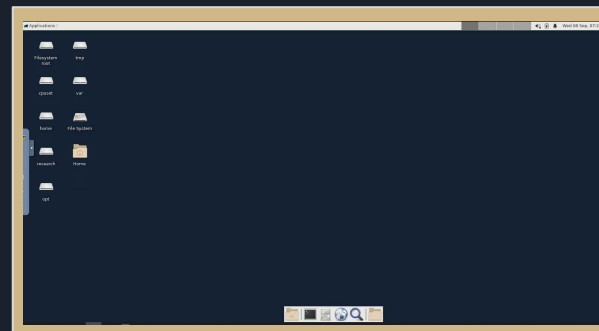
# Ways to Use HPC

## Graphical Interface

- Open OnDemand

## Command-Line Interface

- Any SSH-Supporting CLI Terminal
  - Windows PowerShell
  - Cygwin Emulator
  - MobaXTerm Emulator
  - PuTTY
  - Linux Terminal
  - Mac Terminal



# Open OnDemand

## Easier, more intuitive HPC access

- Desktop interface
- File Transfer

## GUI Applications

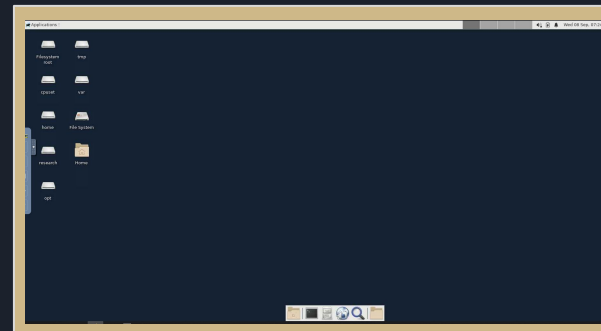
- MATLAB
- STATA
- VisIt

## Servers

- Jupyter Notebooks
- RStudio Server

## Visualizations

- PyMOL
- VMD



[ood.rcc.fsu.edu](https://ood.rcc.fsu.edu)

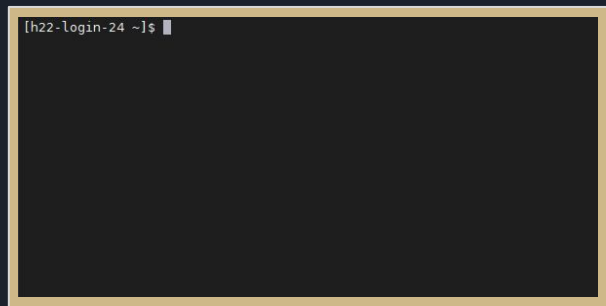
<https://rcc.fsu.edu/docs/ood>

# Command Line Interface

- RCC uses Linux on all of our compute nodes and login nodes
- Command-Line access via SSH (Secure Shell) is one of the primary ways of accessing HPC
  - For Windows, you'll need a terminal emulator like MobaXTerm

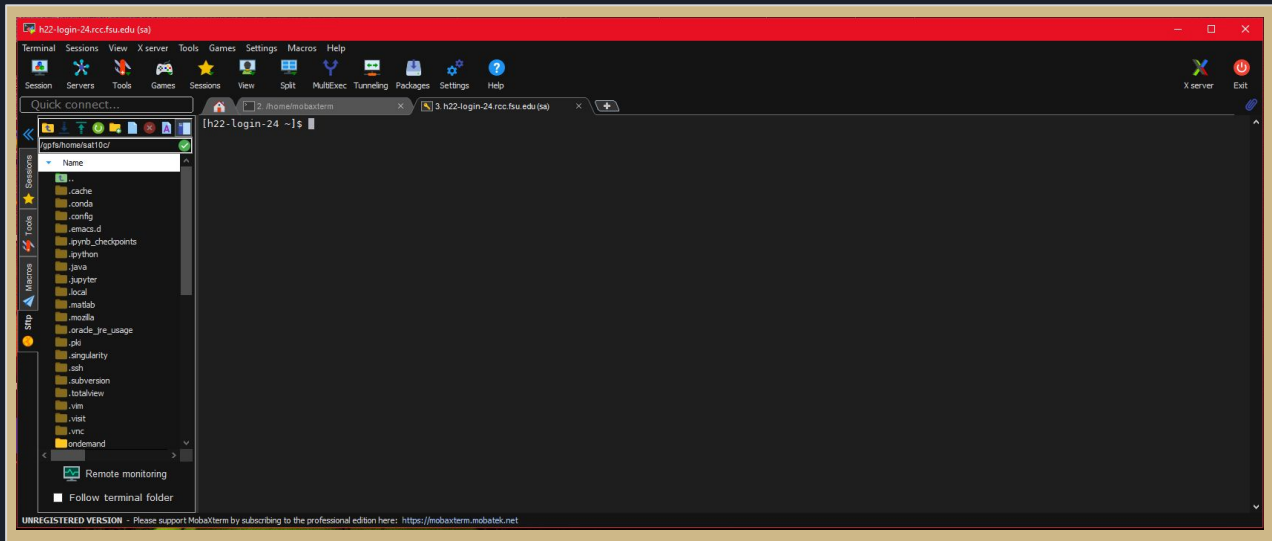
## Tutorials

- [Ryan's Tutorials](#)
- [Ubuntu's Webpage](#)
- RCC also periodically offers an Intro to Linux Workshop



# Mixed CLI/GUI Interface

## MobaXTerm



# Scheduling Jobs

## The SLURM Job Scheduler

- SLURM is the program that finds and allocates resources for you to run your jobs
- Uses a fair share algorithm to help reduce wait times and prevent clogging
  - <https://rcc.fsu.edu/docs/submitting-hpc-jobs-slurm>

## Building Job Scripts

- SLURM is a simple declarative command language built on top of BASH.
- [RCC's Submit Script Generator](#)

```
#!/bin/bash

#SBATCH -J TestProfile
#SBATCH -p backfill2
#SBATCH -n 4
#SBATCH -t 00:10:00
#SBATCH --profile=task

module load intel openmpi

srun trap-intelopenmp1
```

# Non-Interactive Batch Jobs

- Useful for large scale, massively parallel and long-running jobs.
- Set-it-and-forget-it until job is done!

Create a Job Script: "myJob.sh"

```
#!/bin/bash
#SBATCH -J testGROMACS
#SBATCH -n 16
#SBATCH -p backfill
#SBATCH -t 4:00:00

module load gnu openmpi

srun gmx mdrun -v -deffnm em
```

Save file and submit job!

```
[h22-login-24 ~]$ sbatch testscript.sh
```

<https://acct.rcc.fsu.edu/doc/hpc-partitions>

<https://rcc.fsu.edu/docs/hpc-job-reference>

<https://rcc.fsu.edu/docs/hpc-cheat-sheet>



# Interactive Jobs

- Allows direct, interactive access to the actual compute nodes.
- Allows for interactive use of specialized hardware like High-RAM machines, GPU nodes and more.

```
[h22-login-24 ~]$ srun -p backfill2 -t 01:00:00 -n 16 --gres=gpu:2 --pty /bin/bash
```

- **srun** :: “start an interactive job”
- **-p backfill2** :: “... on the ‘backfill2’ queue”
- **-t 01:00:00** :: “...for up to 1 hour”
- **-n 16** :: “...with 16 processors available”
- **--gres=gpu:2** :: “...with 2 GPUs too”
- **--pty /bin/bash** :: “...using the BASH shell”

```
(base) [sat10c@hpc-i36-1 ~]$
```

# Software

## Scientific

Armadillo (C++)  
BOOST  
FFTW  
OpenCV

## Data Science

NiftyReg  
HDF5  
Apache Spark  
R

## Chemistry

ABINIT  
NWChem  
GROMACS  
Wannier90

## Bioinformatics

Trimmomatic  
MIGRATE  
MAUVE  
MAFFT

## Parallel

OpenMPI  
MVAPICH2  
CUDA  
TotalView

<https://rcc.fsu.edu/software>

# Licensed Software

- Typical licensed software installation requires a license acquired by group
  - Access limited to valid license holders to comply with license terms.
- RCC already hosts some licensed software
  - VASP, SPSS, Stata, Agisoft Metashape
- University-wide licenses
  - MATLAB

# Specialized Hardware and Software

## Specialized Hardware

- Contains more processors than the average home computer
- Has several generations of both AMD and Intel Processors
- Can house custom machines (via a research grant) with specialized hardware like special CPUs, DSPs, SSDs, GPUs and otherwise.

## Offload Your Work

- Takes pressure off your regularly used desktop, laptop or other devices
- Scales computational processes not possible on home machines
  - parallel processes
- Runs tasks that need to be repeated many times over a long period of time

# Storage and Support

## Safety

- Store your research data safe from common threats:
  - Disk theft, coffee spills, and regular wear and tear from heavy personal computer usage
- Reduces the risk of data loss

## Data Repositories

- Storage systems which scale into the Petabytes
- Parallel Storage available through GPFS and Archival Storage

## Support

- Workflow process support and assistance with research development
- Software and Systems support ( at [support@rcc.fsu.edu](mailto:support@rcc.fsu.edu) and online documentation)

Questions?

