

UF Research Computing: Environment Modules RHEL6 update Q&A

Matt Gitzendanner: magitz@ufl.edu
Alex Moskalenko: om@hpc.ufl.edu

9/9/12

UF Information Technology www.it.ufl.edu

UF 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

UF Information Technology www.it.ufl.edu

UF Research Computing



UF Information Technology www.it.ufl.edu



UF Research Computing



Quick overview, modules and RHEL6 upgrade

UF Information Technology www.it.ufl.edu


What can you run?


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

UF Information Technology www.it.ufl.edu

Cluster basics

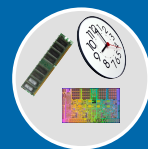
User interaction






Login node
(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

UF 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.

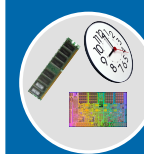
UF Research Computing

- ▶ Submission Script

```
#!/bin/bash
#
#PBS -N My_Job_Name
#PBS -M Joe_Shmoie@ufl.edu
#PBS -m abe
#PBS -o My_Job_Name.log
#PBS -j oe
#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



So what is this “module” thing?

- ▶ **lmod**—Implementation of Environment Modules developed at TACC
- ▶ Allows easy management of user's environment



Lmod: Environmental Modules System

The standard way

```
PATH=$PATH:/some/long/path/to/application
export $PATH
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/long/path/to/place/I/probably/can't/find
export $LD_LIBRARY_PATH
```

- ▶ Need to track down paths to applications, libraries, etc.
- ▶ Multiple compilers, and MPI implementations
- ▶ Manage dependencies
- ▶ Multiple versions of apps



The module way

- ▶ `module load trinity`
- ▶ Automatically:
 - Sets, `$HPC_TRINITY_DIR`
 - To run Inchworm, simply type


```
inchworm --reads reads.fa --run_inchworm [opts]
```
 - Loads Bowtie and Allpaths, two Trinity dependencies
 - You don't need to hunt those down, or worry if they are in your path or not

Module discovery

- ▶ `module spider`
 - List everything
- ▶ `module spider cl`
 - List applications that have cl in name
- ▶ `module spider clustalw/2.1`
 - List details about this version of clustalw
- ▶ `module key molecular`
 - Keyword search for applications

Multiple versions

```
[magitz@submit1 ~]$ module spider gaussian
Rebuilding cache file, please wait ... done
```

 gaussian:

Description:
 A software for electronic structure modeling

Versions:
 gaussian/e01
 gaussian/g03
 gaussian/g09

 To find detailed information about gaussian please enter the full name.
 For example:

```
$ module spider gaussian/g09
```

Multiple variants of a version

```
[magitz@submit1 ~]$ module spider mrbayes/3.2.1
Rebuilding cache file, please wait ... Done
```

 mrbayes: mrbayes/3.2.1

 Description:

Bayesian inference of phylogeny

This module can be loaded directly: module load
 mrbayes/3.2.1

Additional variants of this module can also be loaded
 after the loading the following modules:

intel/2012, openmpi/1.6

Module loading

- ▶ module load raxml
- ▶ module load intel raxml
- ▶ module load intel openmpi raxml
- ▶ module load intel/12 openmpi/1.6
 raxml/3.2

- ▶ module unload raxml

Module swapping

- ▶ module load intel openmpi abyss
 - ▶ module list
- Currently Loaded Modules:
 1) intel/2012 2) openmpi/1.6 3) abyss/default
- ▶ module swap openmpi/1.6 openmpi/1.5.5
- Due to MODULEPATH changes the following modules
 have been reloaded:
 1) abyss



Basic commands

- module spider
- module spider gaussian
- module avail
- module list
- module load clustalw
- module load python/2.6.5
- module add intel openmpi
- module load intel/10.1 openmpi/1.2.7 mrbayes
- module del, rm, unload clustalw - pick one!

Making your own modules

- ▶ You can write module files for your personal applications
- ▶ \$HOME/.modules/app/1.0.lua
 - See Modules page on wiki

RHEL 6 transition

- ▶ Last 4+ years: CentOS 5
- ▶ Transition to RHEL6.3 is **essentially complete**
 - Applications need to be recompiled
 - **File help request** to have applications migrated
 - **You need to update your scripts to use new apps, libraries, etc.**
 - Best done using modules

UF Information Technology www.it.ufl.edu


Common problems

- ▶ Don't use mpi-selector
 - Modules
 - `module load intel openmpi`
 - `module load intel mvapich2`
- ▶ Rebuild your own code in el6
 - Default compiler versions have changed:
 - Intel 2012, GCC 4.4.6, OpenMPI 1.6
 - Library paths have changed
- ▶ Load modules to compile **AND** at runtime
 - To run Intel compiled app add "`module load intel`" as a line in your submit script

UF Information Technology www.it.ufl.edu

Q&A

- ▶ Open to any questions
- ▶ But especially interested in:
 - Where do you need help?
 - What training topics would be relevant?



UF Information Technology www.it.ufl.edu

Training Schedule

- ✓ Aug 28: Intro to UFHPC, getting started
- ✓ Sept 10: Modules, RHEL6 Transition, User Q&A
- ▶ Sept 17: The Linux/Unix Shell - An Introduction
- ▶ Sept 24: Running Jobs, Submission Scripts, Modules
- ▶ Oct 1: Galaxy Overview, The Basics
- ▶ Oct 8: NGS Data Techniques: General Methods and Tools
- ▶ Oct 15: NGS Data Techniques: Reference Based Mapping
- ▶ Oct 22: NGS Data Techniques: de Novo Assembly
- ▶ Oct 29: Phylogenetic Analyses
- ▶ Nov 5: Multiprocessing at the HPC Center
- ▶ Nov 12: Using Git and CMake to Organize and Drive Data Analysis Pipelines
- ▶ Nov 19
- ▶ Nov 29
- ▶ Dec 3

UF Information Technology www.it.ufl.edu