Getting Started on the Clusters

Research Computing Support Group
Division of Information Technology

http://www.rcsg.rice.edu/
rcsg@rice.edu
@ricercsg on Twitter

Chandler Wilkerson
chwilk@rice.edu
Logging in

- Secure shell from within Rice campus
- Use VPN or gw.rcsg.rice.edu for off-campus access
- Linux command-line interface (aka shell)
The Shell

- Provides a prompt
- Interprets input
- Runs commands

[user@host ~]$ ls -la path1 path2...
Environment

- A collection of variables (`env` command)
- Affects the behavior of programs
- Mostly concerned with finding files and describing the system
Finding Installed Software

- Environment Modules
  - module avail
  - module load
  - module list
  - module purge
  - module help
  - env
File Systems

- Everything is a file
- Every file is part of a tree that starts at `/`
- The tree's branches are called directories
- Distinct file systems mount into that tree
Important Cluster File Systems

- Your account (home directory) /home/netid
- Your group (project directory) /projects/groupid
- Shared Scratch (free-for-all) /scratch
- Applications /opt/apps

- Quotas
  - quota -vs
The Job Scheduler

- See it: showq and/or qstat
  - My jobs: showq -u <netid>
  - Queue summary: qstat -q
  - Status of a job: checkjob <jobid>
- Submit jobs: qsub
  - Interactive: qsub -lV -q interactive
  - Batch: qsub filename.pbs
Example Job Script

#PBS -N RJob
#PBS -q serial
#PBS -l nodes=1:ppn=1,walltime=00:10:00
#PBS -M your-email@address
#PBS -m abe
#PBS -V

echo "Job ran on:"
cat $PBS_NODEFILE

R --vanilla < my_R_script.R
Basic R Usage on DAVinCI

- `module load R`
  - Remember `-V` option in `qsub`
- Install R packages in your home:
  - `$ mkdir ~/rlib`
  - `Add the following to ~/.Rprofile`
    - `.libPaths("~/rlib")`
  - `> chooseCRANmirror()`
  - `> install.packages("mypkg", dependencies = TRUE)`