

Using the Linux Cluster at Duke

Tyler Ransom
Duke University

June 5, 2012

Outline

- Linux basics
- Accessing Duke resources remotely
- Submitting computational jobs to the server
- Utilizing text editors

Linux Basics

- Most features of Linux are accessed through the terminal
- This is available through “terminal” on the Linux machines in the Bowling Alley
- On PCs in the bowling alley, use a program like X-Win 32 or SSH Secure Shell Client
- Once logged in, user sees a prompt into which commands can be typed
- Most file-related commands can be executed via the GUI
- All Linux commands have options which are invoked with “-”
- <http://files.fosswire.com/2007/08/fwunixref.pdf> is a good reference for common commands

Accessing Duke resources remotely

- Use a SSH client (PC) or Terminal (Mac or Linux) to remotely access the Duke servers
- The Economics Computing Support (ECS) team has a website outlining how to access various resources
- For help installing SSH software, see this link:
`http://dialog.econ.duke.edu/help/?p=271`
- To remotely view/edit files in your home directory, use an FTP program like FileZilla, WinFTP or Cyberduck

Checking Server Status

- Duke Economics has a batch cluster that allows users to execute computational jobs without being logged in
- You can check how many users are logged in by typing `w`
- `uname -a` tells you which of the three servers you're logged in to (login-02, login-03, login-05)
- `qstat -u "*" displays all active jobs on the cluster`
- `qstat` displays all of your active jobs on the cluster

Submitting a Matlab script to the cluster

- The syntax for submitting a script m-file is `matsub myscript.m myscript.log`
- Once submitted, user gets a message `Your job 12345 ("matsub") has been submitted`
- User must be in the directory where the m-file is located when `matsub` command is issued
- The system creates a file `matsub.o12345` which is a text file containing any errors encountered during execution
- When a program is running perfectly, the `.o` file is useless, but it is very useful when the program encounters errors
- The file `myscript.log` is a text file containing what would have been printed in the command window

Submitting other types of scripts to the cluster

- The syntax for submitting other types of scripts is very similar:
`statasub myfile.do` (no `.log` file needed)
- Users can also submit more general Linux scripts (called shell scripts)
- Computation of lower-level languages like C and Fortran is handled through shell scripts
- Syntax for this is `qsub -q all.q myshell.sh`
- Users can also submit Matlab and Stata jobs via shell scripts if more general options are desired (e.g. have server email you once a job is complete)
- For a complete guide to using the cluster, see
<http://dialog.econ.duke.edu/help/?p=926>