Lecture 2: Useful Tools and Cluster Warm-up (An Interactive Tutorial)

Nicolas Savva

28 Jan 2014
Disclaimer

- Quick overview
  - nothing new for some...
  - overwhelming for others
- Similar information available on class wiki
- Be able to do HW0 with ease after today
- Please ask questions
Logistics

- Everybody in the class should be on CMS by now
  http://cms.csuglab.cornell.edu/web/guest
- The enrollment cap has been increased to the room limit (several seats are available)
- If you cannot see the course under CMS or log in to the cluster please email us with your netID
- HW0 is out (due Tuesday Feb 4th)
Today

- Connecting to the cluster
- Basic usage of the shell
- Version control
- Running and monitoring cluster jobs
- HW0 walkthrough
- More demos (if time permits)
C4 cluster

- Rocks linux cluster (V6.1)
  [http://www.rocksclusters.org/wordpress/](http://www.rocksclusters.org/wordpress/)
- Heterogeneous nodes
- HTCondor scheduler (v7.8.5)
  [http://research.cs.wisc.edu/htcondor/](http://research.cs.wisc.edu/htcondor/)
  use to submit and monitor jobs on cluster
- Ganglia activity monitor
  [http://c4.coecis.cornell.edu/ganglia/](http://c4.coecis.cornell.edu/ganglia/)
C4 cluster machines

- "coecis"
  - 9 nodes (dual Xeon E5345 @ 2.33GHz)

- "compute"
  - 3 nodes (dual Xeon E5345 @ 2.33GHz)

- "cs-instructional"
  - 5 nodes (dual Xeon E5504 @ 2.00GHz)

- "pac"
  - 6 nodes (dual Xeon E5-2690 @ 2.90GHz)

Head node (dual Xeon X5672 @ 3.20GHz)
ssh access to cluster

- Use a Terminal under OS X or Linux
- PuTTY or Cygwin for windows
- Authenticate with Cornell NetID and password
  
  `ssh netID@c4.coecis.cornell.edu`
Terminal prompt

- Type the following for a one-off initialization:

/share/cs-instructional/cs5220/script/setup.sh

This appends commands to .bashrc and .bash_profile to automatically set a class-related environment every time you log in.
Customize the prompt

Google customize $PS1
add your version to the .bashrc file

`export $PS1 = ...`

```
\u username
\h hostname
pwd working directory path
\n new line
```
Some bash commands

ls  w  chmod  ssh
cd  ps  export  scp
pwd jobs set tar
mkdir fg alias vi
rmdir bg exit nano
mv  kill history which
cp  Ctrl + C  |  man
grep Ctrl + Z  find  echo

If you are not particularly familiar with the above please check out the following tutorial:
http://software-carpentry.org/v4/shell/
For more details

- *The UNIX programming environment*
  by Kernighan & Pike
  http://cornell.worldcat.org/title/unix-programming-environment/

- *Learning the bash shell*
  by Newham & Rosenblatt
  http://cornell.worldcat.org/title/learning-the-bash-shell/

Both are available through the Cornell library
Terminal multiplexer (tmux)

- Concurrently view/manage multiple programs in one terminal
- Read documentation (key-binding / configuration)
- Edit .tmux.conf

http://tmux.sourceforge.net/
Version Control Systems (VCS)

- Why version control?
  - Keep revision history
  - Easy way to share files between machines
  - More effective collaboration
  - (Remote) backup
VCS - many flavors

- Distributed
  - Git
  - Mercurial (Hg)
  - Bazaar
  ...

- Client-server
  - Subversion (SVN)
  - CVS
  ...
  ...
Repositories - Host projects online

- Bitbucket (Mercurial or Git)
- Github (Git)
- Launchpad (Bazaar)
- Google Code (SVN, Mercurial or Git)
- Microsoft CodePlex (SVN, Mercurial or Git)
- Sourceforge (CVS, SVN, Bazaar, Git or Mercurial)
- Cornell Forge (SVN) - http://forge.cornell.edu
Class repository and wiki are hosted on Bitbucket:
https://bitbucket.org/dbindel/cs5220-s14/
Sign up for a free account

Git tutorial
https://www.atlassian.com/git/tutorial

Pro Git book
http://git-scm.com/documentation

Git (Interactive) Cheat Sheet
http://ndpsoftware.com/git-cheatsheet.html

Git commands overview
Making a local copy

- Clone the class repository

  ```
  git clone https://bitbucket.org/dbindel/cs5220-s14.git
  ```

  The folder cs5220-s14 now contains local copy
Git basics demo

- `git clone`
- `git add`
- `git diff`
- `git commit`
- `git log`
- `git remote`
- `git pull`
- `git push`

Please look at the cheat sheet and tutorials for more functionality.
Graphical Interfaces available

- **EGit (Git with Eclipse IDE)**
  http://www.eclipse.org/egit/

- **Subclipse (SVN with Eclipse)**
  http://subclipse.tigris.org/

- **SourceTree (Git and Mercurial GUI under Windows or Mac)**
  http://www.sourcetreeapp.com/

- **TortoiseSVN, TortoiseHg, TortoiseGit... series (mostly Windows)**
Eclipse PTP (Parallel Tools Platform)

- IDE for developing parallel applications
- Support MPI, OpenMP, UPC
- Parallel debugger
- Various profiling tools
- Some issues running on c4 cluster
- Experiment with it on your own machine
- Please don’t use PTP remotely on c4 right now

http://www.eclipse.org/ptp/
Using HTCondor

- condor_status [-claimed -avail -master -verbose]
- condor_q [ -analyze -run -hold]
- condor_submit
- condor_hold
- condo_release
- condor_rm
- condor_history [ -backwards -forwards -match]

[ -constraint -format ]
HTCondor CS5220 wrapper scripts

- csub (serial submissions)
- mpisub (Message Passing Interface)
- upcsu (Unified Parallel C)
- ompsub (OpenMP: Open Multi-Processing)

https://bitbucket.org/dbindel/c4-pkg
Environment modules

Make it easy to install:

- different versions of software packages
- software packages that might conflict

- module list
- module avail
- module load
  module add
- module unload
  module rm

https://bitbucket.org/dbindel/cs5220-s14/wiki/modules
Homework 0 walkthrough

Steps:

▶ Clone class repo

▶ membench

▶ membench with ClassAd requirements

▶ HTCondor job management

▶ pinfo

▶ Retrieve the results (using sftp or scp)

http://bitbucket.org/dbindel/cs5220-s14/wiki/HW0
Additional Demos...