Cluster Orientation

Henry Linder
m.henry.linder@uconn.edu

September 12, 2016

- Documentation and examples are availabe on the cluster wiki.
- Access the cluster with SSH at `stats.phys.uconn.edu`
  - Windows: Putty
  - OSX, Linux: `ssh` command in terminal
    ```
    $ ssh hlinder@stats.phys.uconn.edu
    ```
- Copy files to / from the cluster with SCP (secure copy)
  - Windows: WinSCP
  - OSX: Transmit
  - OSX, Linux: `scp` command
    ```
    $ scp file hlinder@stats.phys.uconn.edu:~/
    ```
- The cluster can only be accessed on campus. From off campus, do one of the following:
  - Use the UConn VPN
  - Use Skybox
  - SSH into one of the research lab Linux computers, and then SSH again into the cluster (ask Professor Yan for an account)

Makefile

```
setup:
        mkdir err log out rdata
submit:
        condor_submit demo.condor
check:
        condor_q hlinder
data:
        mv *.Rdata rdata/
        Rscript combine.R
clean:
        rm err/* log/* out/* rdata/* *.Rdata
reset:
        rm -r err log out rdata
```

demo.condor

```
universe = vanilla
requirements = ParallelSchedulingGroup == "stats group"

# Code to run on the cluster; make sure it is executable by
    running
# $ chmod +x demo.R
# from the commandline.
executable = demo.R
# Commandline arguments to pass to code
arguments = $(process)

# Log file
log = log/demo.log.$(process)
# Error file
error = err/demo.err.$(process)
# Capture terminal output
output = out/demo.out.$(process)
```

demo.condor

```
# Transfer any new files
should_transfer_files   = YES
when_to_transfer_output = ON_EXIT

# By default, each all new files created for a job are
    copied back to
# the main job directory.
#
# But, it's also possible to explicitly specify which
    files to copy
# back:
# transfer_output_files =

# How many jobs to submit
queue 10
```

demo.R

```
#!/home/statsadmin/R/bin/Rscript

## Read process number
args <- commandArgs(trailingOnly = TRUE)
proc.no <- as.numeric(args[1])

print(paste("Executing process number", proc.no))

x <- rnorm(20)
save(x, file = sprintf("rnorm%02d.Rdata", proc.no))

Sys.sleep(30)
```

combine.R

```
datadir <- "rdata/"

all_x <- c()
for (f in list.files(datadir)) {
    load(file.path(datadir, f))
    all_x <- c(all_x, x)
}

save(all_x, file = "rnorm.Rdata")
```