High Performance Computing

James Rosindell

28/11/2017

Look up how to log in to HPC
HPC is useful for embarrassingly parallel problems (graphics, simulations with multiple parameters) and Non embarrassingly parallel problems (fluid dynamics, a lot of the tasks run by a single problem).

1 How do you parallelize your code?

Imagine before using HPC you want to do one simulation then 3 parallel in HPC.
Important line of code:
$as.numeric(Sys.getenv("PBS_ARRAY_INDEX"))$ Using PC: for i in 1:10 do simulation(i)
$HPC : i <- as.numeric(Sys.getenv("PBS_ARRAY_INDEX"))do_simulation(i)$
$do_simulation <- function(i)setrandomseedasi, selectyoursimulationparameters, doyoursimulat$
$Usedifferentrandomnumberseedswhenrunningparalleltoensurethattheyareindependent.$
$Handlingmemory : Saveyourresultsinmemoryandthenwritetodiskattheend.Outputyourcodetoase$

2 Running code on a cluster

Will be using: cx1 - a cluster of many ordinary computers. Access via a login node. Login node, your potal to cx1, Login.cx1.hpc.ic.ac.uk
Where is data stored?
$Home- backed up main area 10GB
$Work - for running jobs not backed up 150GB
$TMPDIR - Ignore these temporary files but be careful you don't lose work in here
Need to write command in shell script to move files back from tmpdir to home.
Step1: Go to filezilla-project.org

Download FileZilla Client and install it
Open FileZilla and put in the following settings:
(see dropbox)