

# D-place FARM documentation: Running simulations on the cluster

*Ty Tuff, Bruno Vilela, and Carlos Botero*

*project began: 15 May 2016, document updated: 04 July 2017*

Running an R script on the cluster requires two parts: an R script with the code to be run and a PBS script to control how that R script is run on the cluster.

There are two different clusters at Wustl.edu, an old and a new cluster. This script runs the R code on the new cluster.

```
#!/bin/bash
#PBS -N Four_model_run
#PBS -V
#PBS -l walltime=23:59:00
#PBS -l pmem=1200mb
#PBS -l nodes=1:ppn=1:haswell
#PBS -t 1-1000

echo $PBS_ARRAYID

cd /home/cbotero/mydirectory/Four_model_compare
module load R

export R_LIBS=$HOME/rlibs
#R CMD INSTALL --library=/home/ttuff/rlibs FARM_1.0.tar.gz

Rscript --vanilla ./FARM_four_model_compare.R ${PBS_ARRAYID}
```

We were regularly causing problems running to many jobs on the new cluster and we were asked to move to the old cluster. This cluster has slower individual processors, but we can run more jobs at one time, so productivity has stayed about the same.

```
#!/bin/bash
#PBS -N FARM_third_run_old
#PBS -V
#PBS -l walltime=160:00:00
#PBS -l pmem=1200mb
#PBS -q old
#PBS -l nodes=1:ppn=1:nehalem
#PBS -t 1-500

echo $PBS_ARRAYID

cd /home/ttuff/mydirectory/Four_model_compare
module load R

export R_LIBS=$HOME/rlibs
#R CMD INSTALL --library=/home/ttuff/rlibs FARM_1.0.tar.gz
```

```
Rscript --vanilla ./FARM_four_model_compare.R ${PBS_ARRAYID}
```

The final argument in the #PBS script above (#PBS -t 1-500) controls the serial running schema for running many simultaneous instances of the R script at a time. This argument is passed to R as an integer value using the following arguments inside R.

```
args <- commandArgs(trailingOnly = FALSE) #7 elements are passed from the PBS  
NAI <- as.numeric(args[7]) # the seventh of those elements is the array integer.
```

Here is a working example of how to set up the R script.

```
#install.packages("rfoaas")  
library(rfoaas)  
  
##If the PBS script started this code running and passed the number 13  
##to this particular run of a larger serial set.  
  
#args <- commandArgs(trailingOnly = FALSE)  
NAI <- 13 #as.numeric(args[7])  
  
sayHello <- function(loop_number){  
  print(paste0("I can count to ", loop_number, "! ", cool(from="Ty")))  
}  
  
sayHello(NAI)  
  
## [1] "I can count to 13!   Cool story, bro. - Ty"
```

You logon to the cluster using linux/unix code from the command line terminal on your computer. Open the terminal and put in your login info.

```
ssh -Y ttuff@login.chpc.wustl.edu  
password:_____
```

Upon first login, you will be in a folder called 'HOME' with a series of system files in it. You will want to use an FTP client to view and organize these files. I prefer Filezilla, but there are several other good clients available for free. Download filezilla, make sure it's in your applications folder, and open it. You should see a window that looks like a newer version of this. The left panels are the files on the computer you're working from and the right two panels will show the files on the server once you log in through Filezilla also.

Once logged in, you need to change the directory

```
cd /home/ttuff/mydirectory/Four_model_compare
```

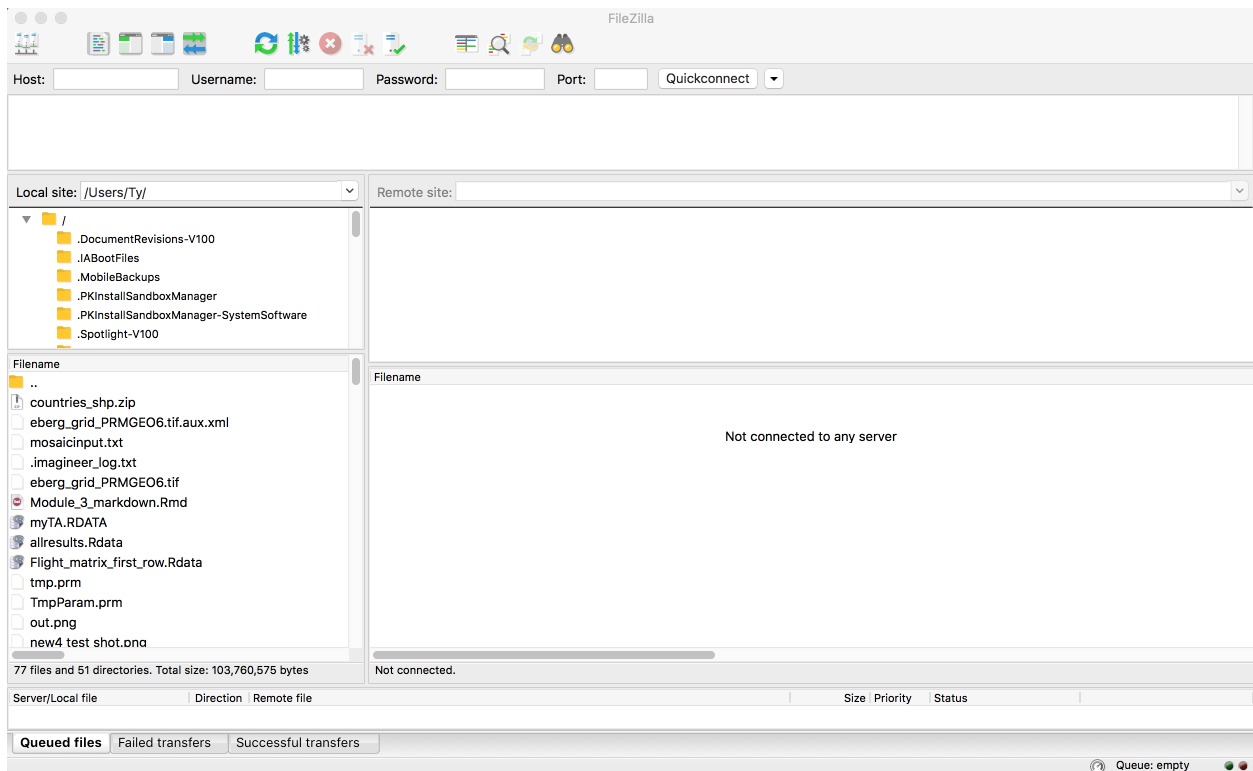


Figure 1: A fresh Filezilla window

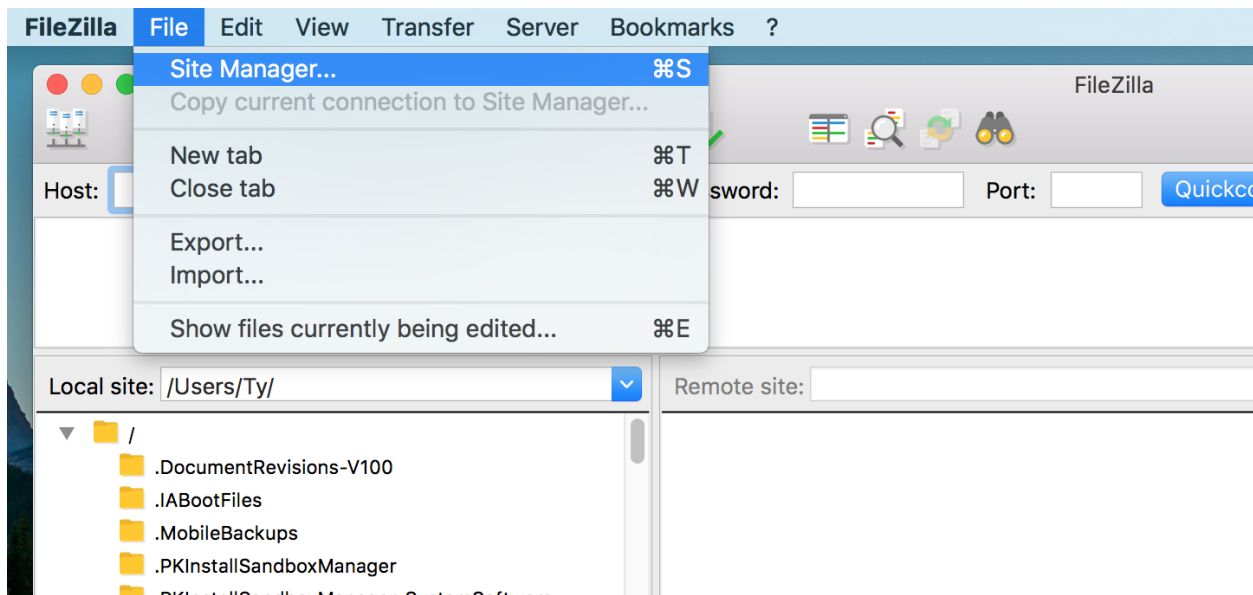


Figure 2: Start a new server link

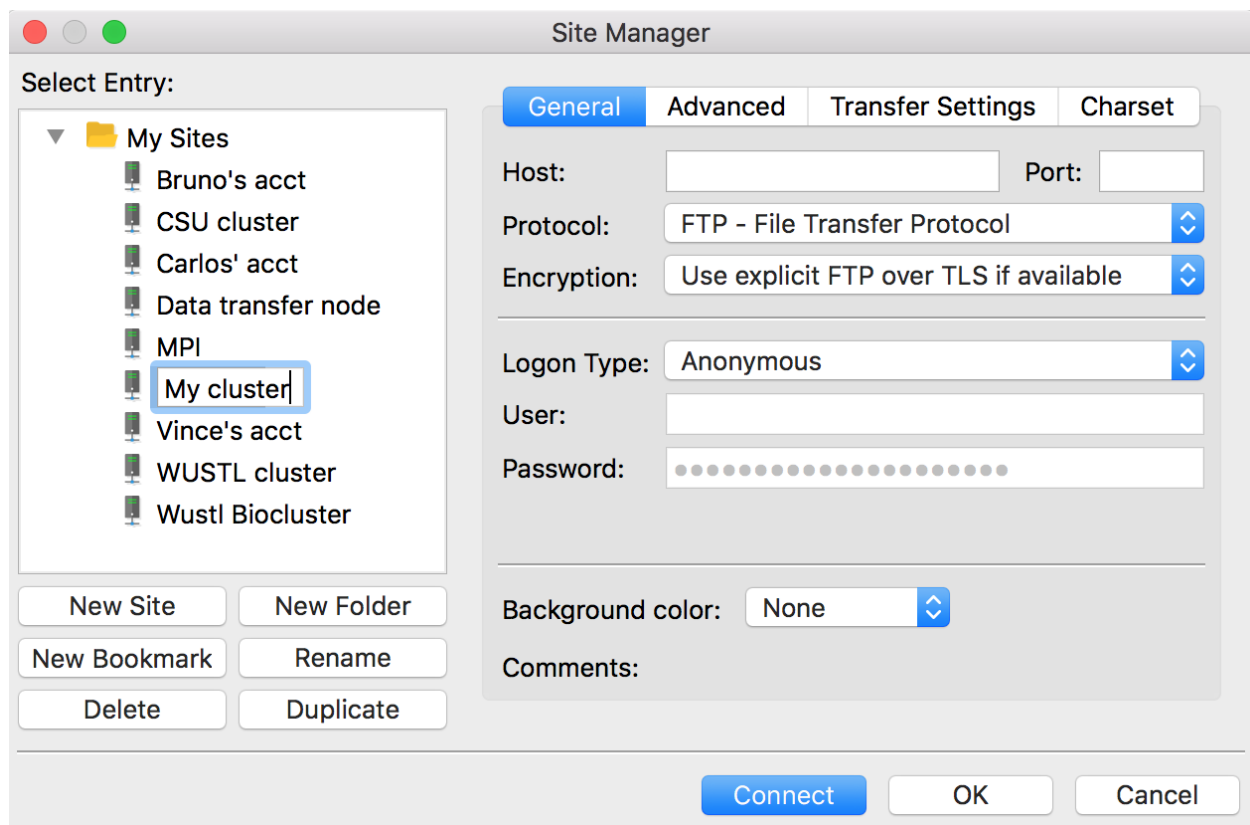


Figure 3: Start a new site and name that new site

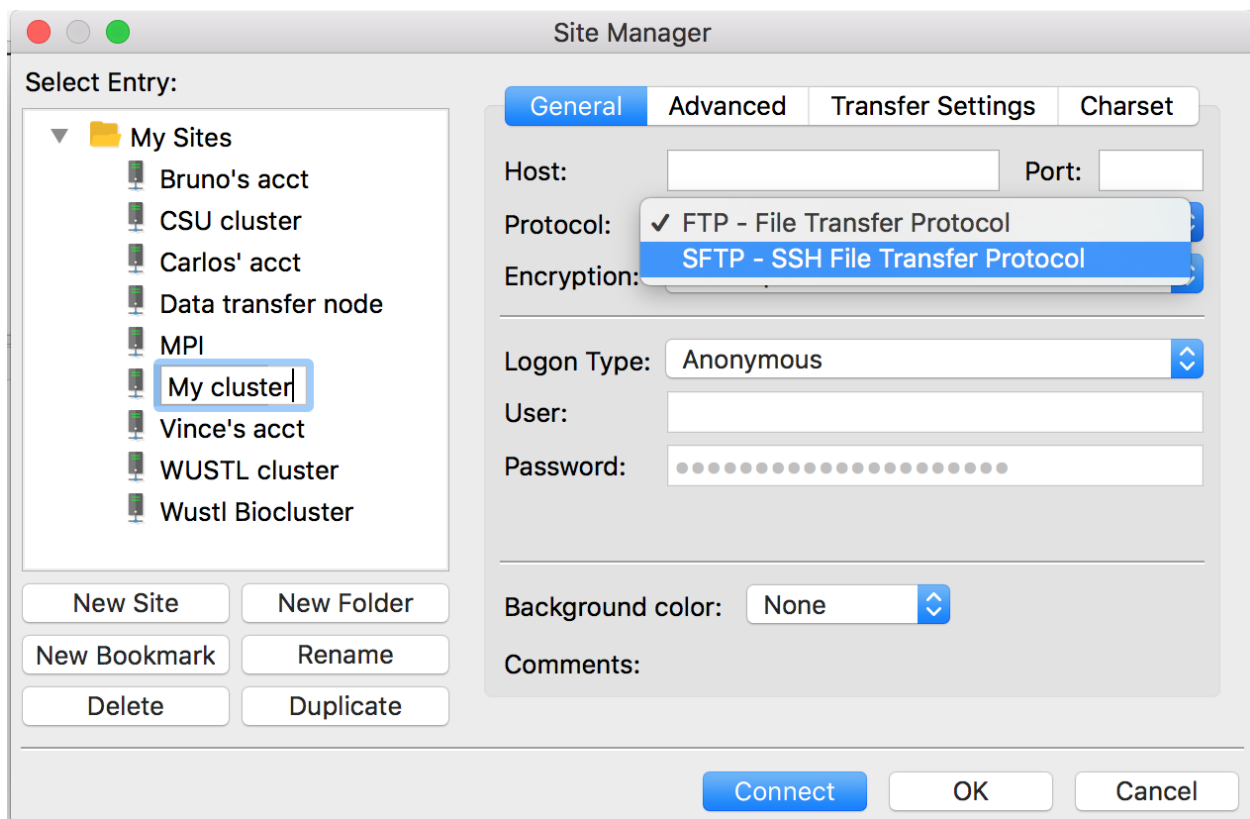


Figure 4: Select a secure ssh file transfer protocol

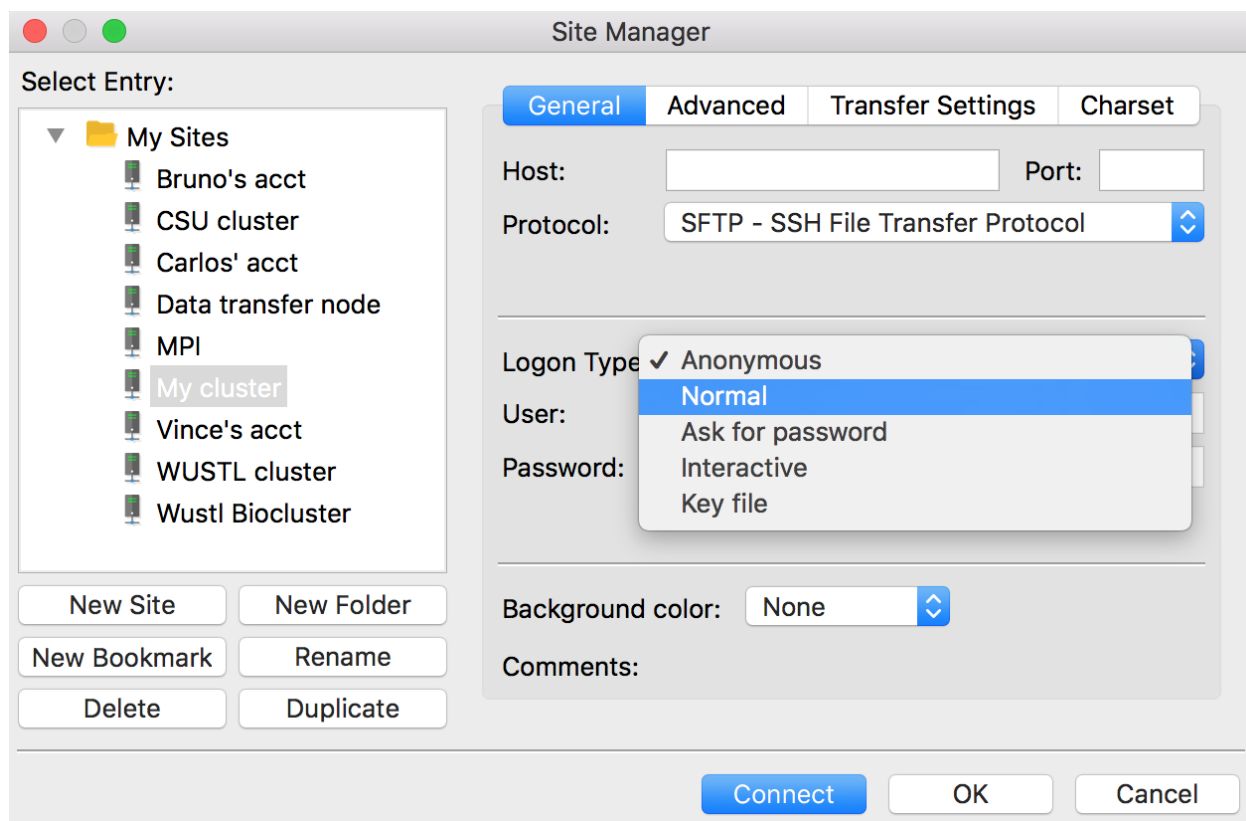


Figure 5: Login as normal

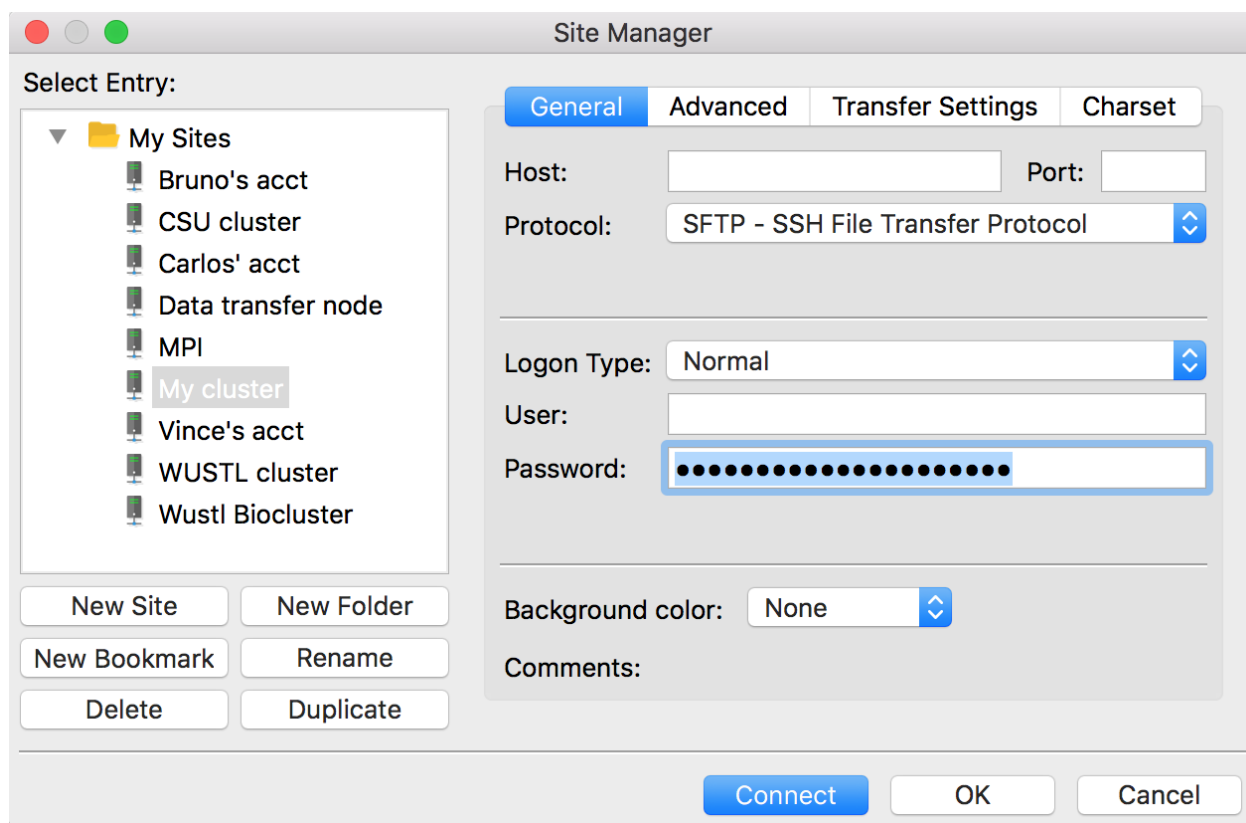


Figure 6: Enter password

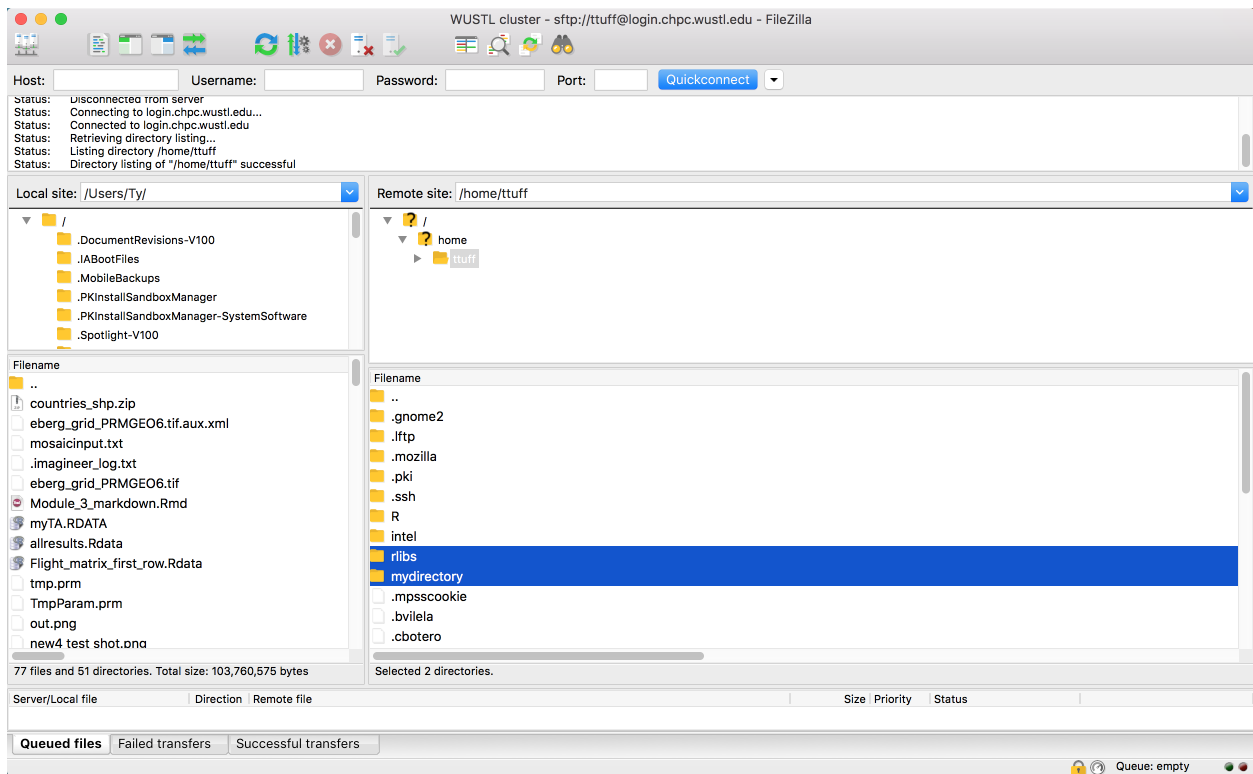


Figure 7: You need to create two new directories (folders) for us to work out of. Call one 'ribs' and the other 'mydirectory'

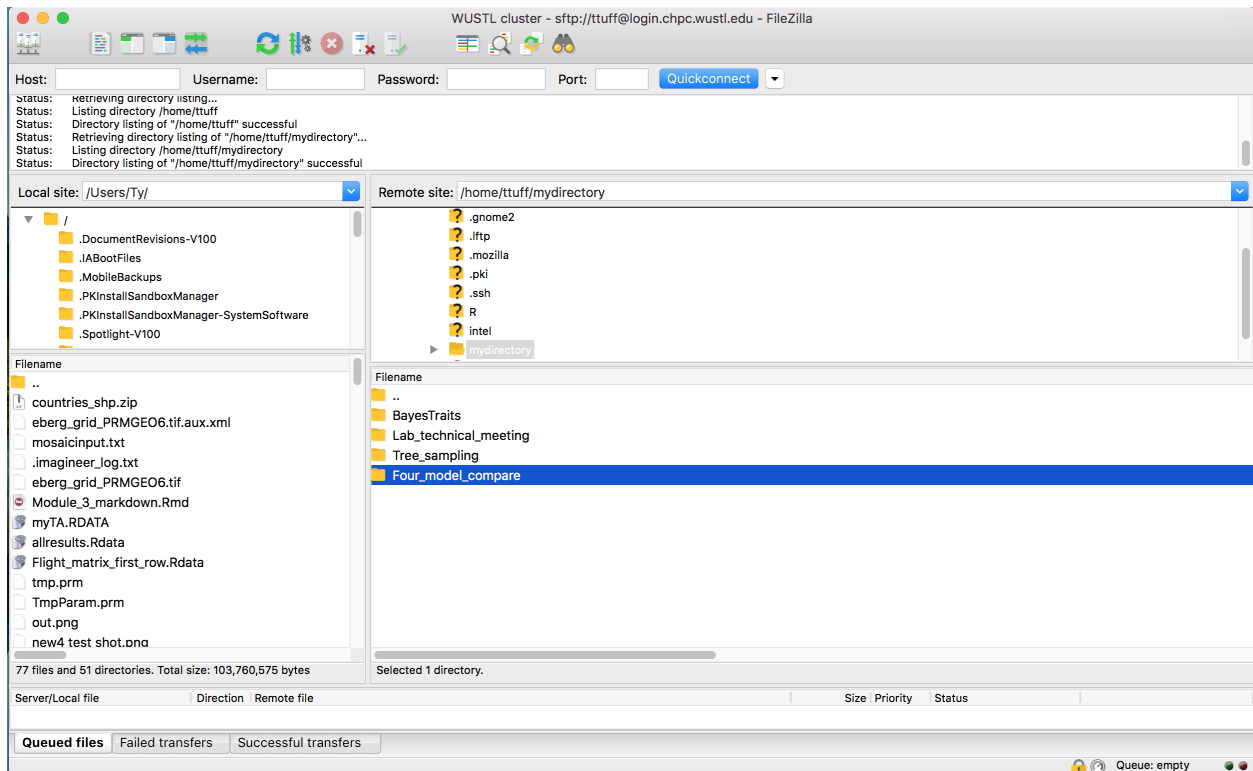


Figure 8: Within mydirectory, create another folder called 'Four\_model\_compare'



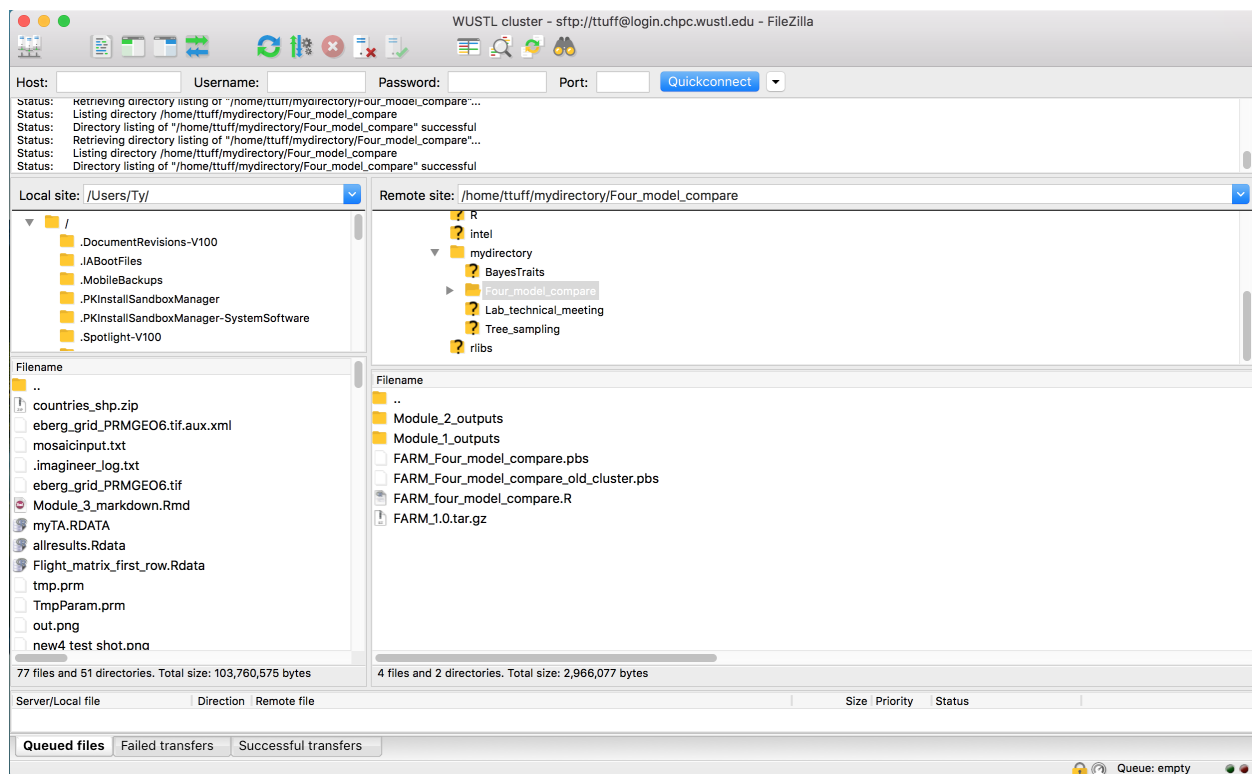


Figure 9: Within ‘Four\_model\_compare’, create two folders called ‘Module\_1\_outputs’ and ‘Module\_2\_outputs’. Then drag three files from your computer files on the left to the server folders on the right: one .R file, one .pbs file, and a zip file with the package FARM in it. These should be named FARM\_four\_model\_compare.R, FARM\_four\_model\_compare\_old\_cluster.pbs, and FARM\_1.0.tar.gz .

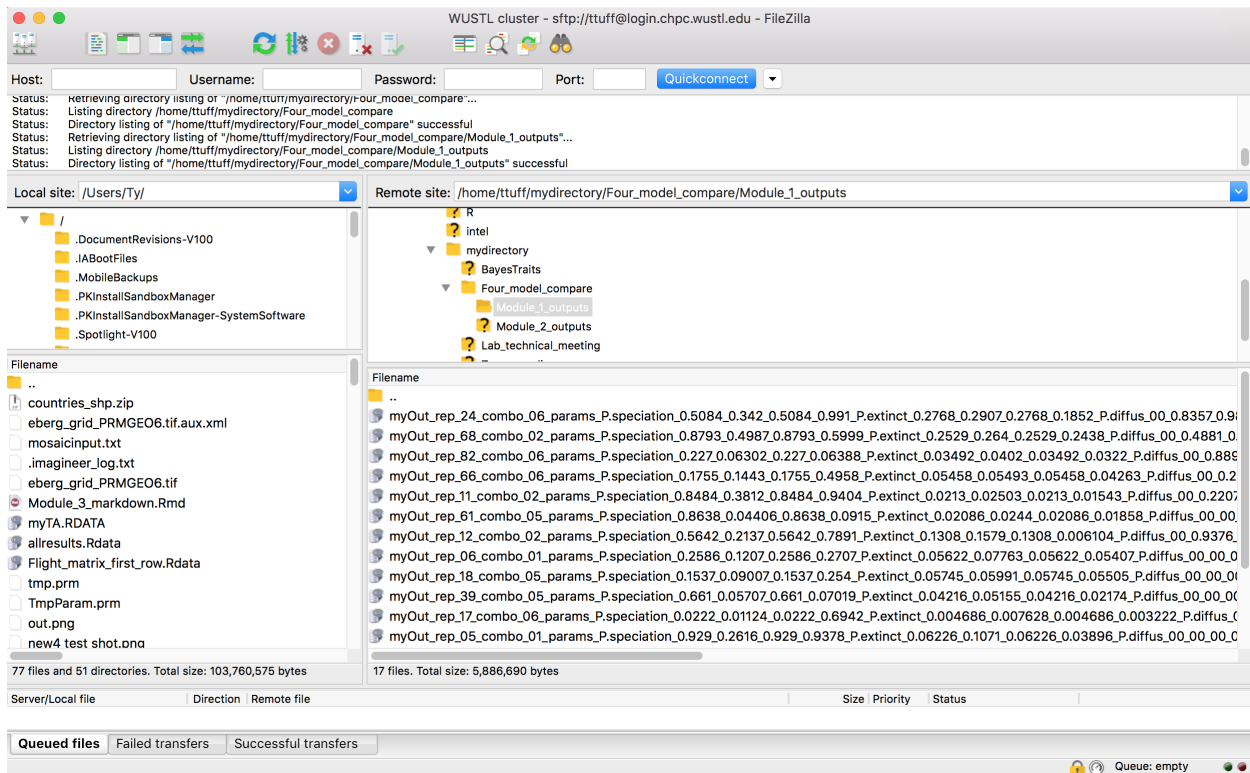


Figure 10:

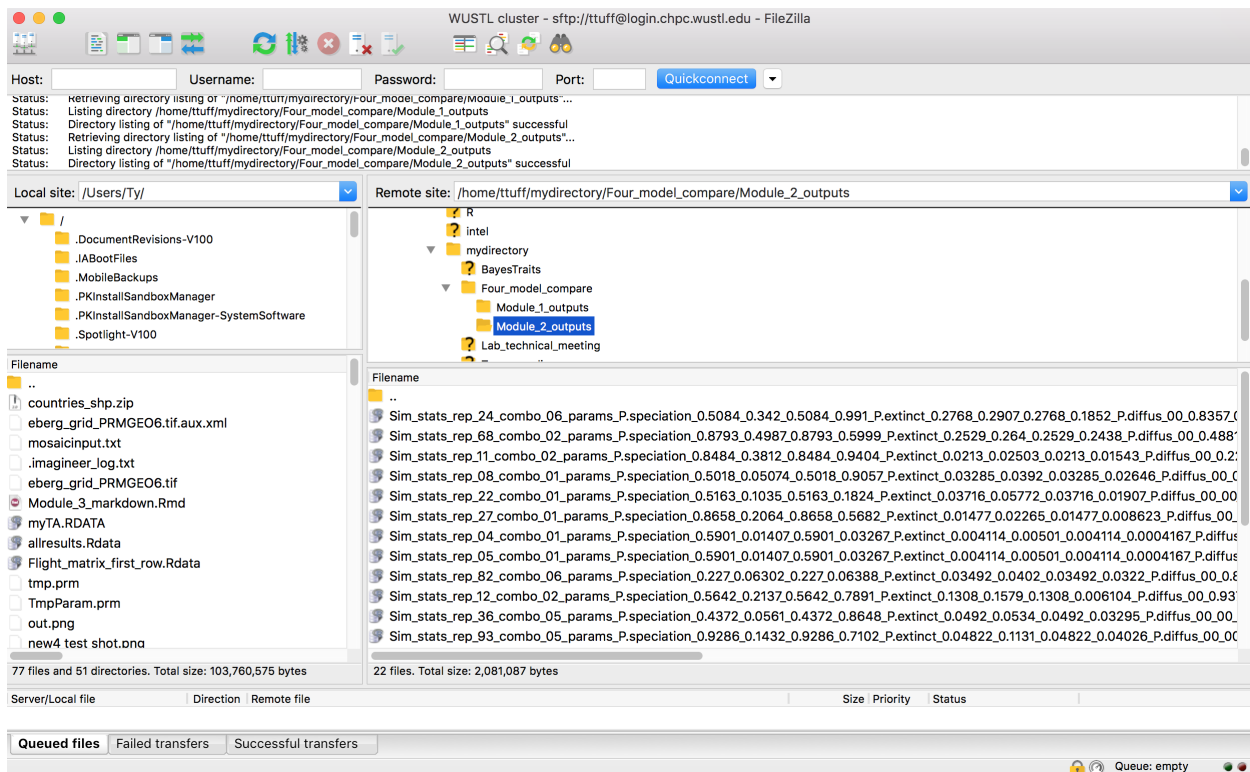


Figure 11: