

Simple batchtools example with SSH

Summary: to make batchtools work with SSH we need to

1. Set up passwordless ssh from the head node to the compute node(s)
2. Install and configure NFS
3. Gloss over one (or two?) small bug. Install from github sampoll/batchtools

1. Setup

Use Bioconductor 3.7 AMI: ami-2951fa53

Launch two EC2 t2.small instances.

Arbitrary choice:

head node: (public) 184.72.186.175 (private) 172.31.50.47

compute node: (public) 54.92.240.136 (private) 172.31.48.86

```
1 R> install.packages("devtools", repos = "http://cran.us.r-project.org")
2 R> library(devtools)
3 R> install_github("sampoll/batchtools")
```

2. Passwordless ssh

1. generate key pair on head node, empty password (`ssh-keygen -t rsa`)
2. use local machine to sftp public key from head node to compute node
3. append to `.ssh/authorized_keys` on compute node
4. on head node, add line to `/etc/hosts` (`54.92.240.136 ec2-54-92-240-136.compute-1.amazonaws.com`)

3. Install and configure NFS

```
1 # on head node (184.72.186.175)
2 sudo apt-get update
3 sudo apt-get install nfs-kernel-server
4 sudo mkdir -p /scratch
5 sudo chown nobody:nogroup /scratch
6 sudo chmod -R 777 /scratch
7 sudo vim /etc/exports
```

```

8 # /scratch 54.92.240.136(rw, sync, no_root_squash, no_subtree_check)
9 sudo systemctl restart nfs-kernel-server
10 # open port 2049 in security group
11
12 # on compute node (54.92.240.136)
13 sudo apt-get update
14 sudo apt-get install nfs-common
15 sudo mkdir -p /scratch
16 sudo chmod -R 777 /scratch
17 # open port 2049 in security group
18 sudo mount 184.72.186.175:/scratch /scratch

```

4. Run example

On master node, cd to /scratch:

```

1 # batchtools.conf.R
2 workers <- list(Worker$new("ec2-54-92-240-136.compute-1.amazonaws.com", ncpus=1))
3 cluster.functions = makeClusterFunctionsSSH(workers)

1 # piApprox.R
2 piApprox = function(n) {
3   nums = matrix(runif(2 * n), ncol = 2)    # random point in [0,1] x [0,1]
4   d = sqrt(nums[,1]^2 + nums[,2]^2)        # distance from origin to point
5   4 * mean(d <= 1)                         # count points in unit circle
6 }

1 R> source("piApprox.R")
2 R> library(batchtools)
3 R> reg <- makeRegistry()
4 R> batchMap(fun = piApprox, n = rep(1e6, 3))
5 R> submitJobs()
6 R> waitForJobs()
7 R> reduceResults(function(x,y) x+y)/3      # 3.1398

```