Account ID: 347645305145
pragati.mitra@fuw.edu.pl

Notes for AWS pcluster:

This is a compilation of notes relevant for launching a parallel cluster (pcluster) in AWS cloud. The tutorial https://www.hpcworkshops.com/ is a very good starting point to learn and launch a pcluster from scratch using AWS CLI.

Two main steps : 1. To create a key-pair, this will be needed to launch and later to ssh/scp to/from the cluster. 2. Launch the cluster with a config file (see example file *.yaml) that contains relevant details to setup the cluster: no of nodes, type of machine, storage memory etc etc…

Command to create key:
aws ec2 create-key-pair --keyname hpckey --query KeyMaterial --output text > hpckey

Command to create pcluster with the config file:
pcluster create-cluster --cluster-name hpc-zhaires-large --cluster-configuration config_zhaires.yaml

Login to headnode from cloud9:
pcluster ssh --cluster-name hpctest2 -i mykey

(upgrade the cluster, remember it does have python but not numpy,scipy etc. Be sure to install those)

Submitting a job
Sbatch job.sh -N 3 -p spot/ondemand

Get exact ip address to ssh to/from:

> In the AWS EC2 console, click on the checkbox next to your instance's name, then click on `Actions` and select `Connect`. Click on the `SSH client` tab and copy the ssh command example.

Copy from pcluster to local:

Command example:

scp -r -i mykey
ec2-user@ec2-52-58-54-78.eu-central-1.compute.amazonaws.com:/home/ec2-user/aires/cluster_stshp/StShp_XD_XD_EPLHC_Proton_2.0_38.2_180.0_1 /home/pragati/Documents/data/

Better: rsync
rsync -avzh -i mykey
ec2-user@ec2-3-65-39-125.eu-central-1.compute.amazonaws.com:/shared/donejobs/*
/media/pragati/

If the above does not work, try only with headnode ip address: (example)ssh -i mykey.pem ec2-user@ip-172-31-16-120:/home/ec2-user/