— Common steps — # 3.3 Setting up ssh and pdsh

Although ssh is not necessary for standalone operation, it is required for pseudo-distributed operation and a cluster setup.

1. Check if ssh (ssh remote login client) and pdsh (issues commands to group of hosts in parallel) is installed > `ssh -V && pdsh -V  //Check if versions are returned`
2. If its not installed, install ssh and pdsh > `sudo apt-get install ssh` > `sudo apt-get install pdsh`
3. Verify if ssh server is running: > `which sshd  //active if it returns path of ssh daemon` > **OR** > `sudo systemctl status ssh //active if it returns "ssh.service: ... active"`
4. To check if that worked, try ssh-ing into localhost > `ssh localhost` > > You should get the following: > `...  Are you sure you want to continue connecting (yes/no)? yes  ...  Welcome to Ubuntu 20.04 LTS ...  ...`
5. Exit this superfluous self-connection by executing the following command > `exit`
6. Change rcmd type of pdsh connection from rsh to ssh > `export PDSH_RCMD_TYPE=ssh` > `pdsh -q -w localhost  //to verify if rcmd_type is ssh` > > Or you could just add that statement to *~/.bashrc* to set RCMD type as ssh everytime you open the terminal. To add it, execute the following command: > `echo 'export PDSH_RCMD_TYPE=ssh' >> ~/.bashrc`
7. Create a rsa public-private keypair without a passphrase in the default location (if you haven't already) > **Either** execute the following command and repeatedly hit enter until everything is done > `ssh-keygen` > > **OR** execute the following: > `ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa`
8. Append public key of generated pair to authorized keys and set user permissions read-only > `cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys` > `chmod 0600 ~/.ssh/authorized_keys`

Sources

1. Apache Hadoop Documentation (Main)
2. Dev Tutorial (Main)
3. Introduction to public-key cryptography
4. Why should I change RCMD type?