Hadoop notes

Hadoop installation steps

Install VMWare workstation (or equivalent) and start it

Create new virtual machine:

Select typical installation option

Select option I will install operating system later

Select a name for virtual machine

Selected 20gb as maximum disk capacity

Select option store virtual disk as a single file

Incrased ram allocation to this vm (8gb )

Selected cd/dvd option to point to ubuntu server iso image file which was downloaded into your machine

Click finish

Above created a blank virtual machine.

3. Install Ubuntu server in this virtual machine

1 power on the virtual machine

Import screens men options during ubuntu installation process

1. Do not select keyboard detection
2. Host name select “Hadoop” as the host machine
3. User name select “Hadoop” as the user name
4. Password select “123456” as password
5. DO NOT ENCCRYPT your home directory
6. Partitioning the disk select guided disk option first option
7. Select save changes to the disk Yes
8. HTTP proxy information leave it blank
9. Managing upgrades not automatic
10. Use space bar to select ssh server (first option)
11. Install grub boot loader …Yes

Now that complete log in with username Hadoop and password 123456

Check the ipaddress of your new VM

Use command ifconfig….192.168.11.128

4. Update os

Sudo apt-get update

5. install OpenSSH Server

Sudo

8. Extract the tar file

Tar -xvzf Hadoop-1.2.1.tar.gz

9. Copy the contents of the extracted folder Hadoop 1.2.1 into the location /usr/local/Hadoop

Directory conventions

Windows Linux

Root dir C:\ /

Software c:\Program files /bin

Framework c\windows\System32 /usr/local

Sudo cp -4 hadoop-1.2.1 /usr/local/Hadoop

10. inform system (Linux)wher is Hadoop

Sudo vi /home/Hadoop/.bashrc

{to edit the file in vi editor… press insert}

After fi add:

export HADO\_PREFIX=usr/local/Hadoop

export PATH=$PATH:$HADOOP\_PREFIX/bin

Use insert to get into edit mode

escape makes

to save the file and quit escape and :wq

sudo nano /home/Hadoop/.bashrc

(go do bottom and add exports at bottom and use menu options to save and escape)

11. Update the bash settings

Exec bash

1. Inform Hadoop where is java installed on your system

In winscp:

Go to usr.local Hadoop

Look in bin for binary files

Lib folders contains all the jar files

Conf configuration files to configure Hadoop environment

Hadoop-env-sh (click on it)

Set environment variables in it

Set JAVA-HOME variable to where you have set up java.

Open editor in linux to change Hadoop-env.sh file

Sudo vi /user/local/Hadoop/conf/Hadoop-env.sh

(uncomment the java\_home variable and change to):

Export JAVA\_HOME=/user/lib/jvm/java-1.7.0-openjdk-i386

(ensure this path is current)

Export JAVA\_HOME=usr/lib/jvm/java-1.7.0-openjdk-i385

{to save and quit esc :wq }

At this stage you have successfully installed Hadoop in standalone mode

Type Hadoop to get help commands to know if you have installed correctly

Ubuntu 16.04 to use jdk 8