

INSTALLATION OF HADOOP IN UBUNTU USING VIRTUALBOX

• INSTALL OF VIRTUAL BOX

- To download virtual box, go to official site "virtualbox.org"

- Download the latest version for windows

STEPS:

- Getting started in virtual box 6.1.2. setup wizard click Next

- Select Installation location - location can be user's choice then click Next

- Creating entries and shortcuts. in custom setup check all the boxes and then click Next

- Ready to install - After this ready to install box will appear, click on install button

- After this all files and packages will be installed

- After successful installation, in windows security box, click the install button

- After this click Finish button

- Virtual Box is successfully installed

• INSTALL UBUNTU ON VIRTUAL MACHINE

PRE REQUISITE

Should have installed Virtual box or other hypervisor

- Download Ubuntu AND 64 bit ISO file version 22.04 from ubuntu official website
- Open Virtual Box and click New Button
- Select the name and location of where you want to install your ubuntu machine
- Allocate hard disk size, RAM & etc to your virtual machine click on Next
- After disk allocation is done, provide the boot and begin installing Ubuntu
- Browse the file location and select the ISO file for Ubuntu
- Proceed and Ubuntu will be installed
- Choose a name, set password to login and click continue
- Installation process is over then boot the virtual machine.

◦ INSTALL HADOOP IN VIRTUAL BOX

- Prepare the environment
- Install Java 8 in your computer
- Create a hadoop user
- Install ssh for hadoop
- Create a setup ssh key based authentication
- Download and install hadoop
- Configure environmental variables
- Initialize hadoop distributed file system

Other process

- Setting up non-root user for hadoop environment :
sudo adduser hadoop:
m - hadoop
- Enabling passwordless ssh for the hadoop user
ssh-keygen -t rsa -p
- Editing configuration files such as
hadoop-env.sh, core-site.xml, hdfs-site.xml,
mapred-site.xml, yarn-site.xml
- Formatting HDFS namenode
- Starting the hadoop cluster
- Accessing the hadoop UI from browser
- Verify installation : jps
- Setup environmental variable
source ~/.bashrc
echo 'export PATH=\$PATH:/usr/local/hadoop/bin:/usr/
local/hadoop/<bin>' ~/.bashrc.