Hadoop Linux - Pseudo Distribution

Download Ubuntu 16: <https://www.ubuntu.com/download/desktop>

Installing Hadoop Step by Step:

<http://www.bogotobogo.com/Hadoop/BigData_hadoop_Install_on_ubuntu_single_node_cluster.php>

**Note**: Above link is missing YARN configuration

Refer following link for YARN configuration

<https://www.ibm.com/developerworks/community/blogs/d9a07ec3-11e2-467d-b758-6861c4cb1d44/entry/How_to_install_Hadoop_2_7_0_in_ubuntu_16_04?lang=en>

Hadoop Windows - Pseudo Distribution

1. Extract hadoop-2.6.2 file on your local system where you want to install Hadoop
2. Copy java JDK in C: drive if already not there
3. Go to location **hadoop-2.6.2/etc/hadoop** 
   1. Modify core-site.xml file to update hadoop.tmp.dir property

Create temp directory somewhere and use that as value to mentioned property

* 1. Modify Hadoop-env.cmd to update JAVA\_HOME

Should be the location of your JDK in C: directory

* 1. Hdfs-site.xml:

Create directories /hadoop-2.6.2/data/namenode and /hadoop-2.6.2/data/datanode

Update dfs.namenode.name.dir and dfs.datanode.name.dir properties accordingly

* 1. yarn-site.xml

Update yarn.nodemanager.log-dirs property

1. set HADOOP\_HOME environment variable to bin directory within hadoop-2.6.2 folder
2. open windows command prompt cmd and execute following commands
   1. hdfs namenode –format
   2. change directory to hadoop-2.6.2/sbin
   3. start-dfs.cmd
   4. start-yarn.cmd

# Cloudera Quick Start VM:

<https://www.cloudera.com/downloads/quickstart_vms/5-12.html>

Select VMWare as platform