1. Recommended Platform

OS : ubuntu 12.04

Hadoop: Apache Hadoop(hadoop-2.2.0 or above)

1. Install java

2.1) Update the source list

command:-sudo apt-get update

2.2) install open jdk

command:-sudo apt-get install openjdk-6-jdk

1. Configure SSH (shown below in screenshot):-

3.1) install open SSH Server-Client

command:-sudo apt-get install openssh-server openssh-client

3.2) Generate key value pair

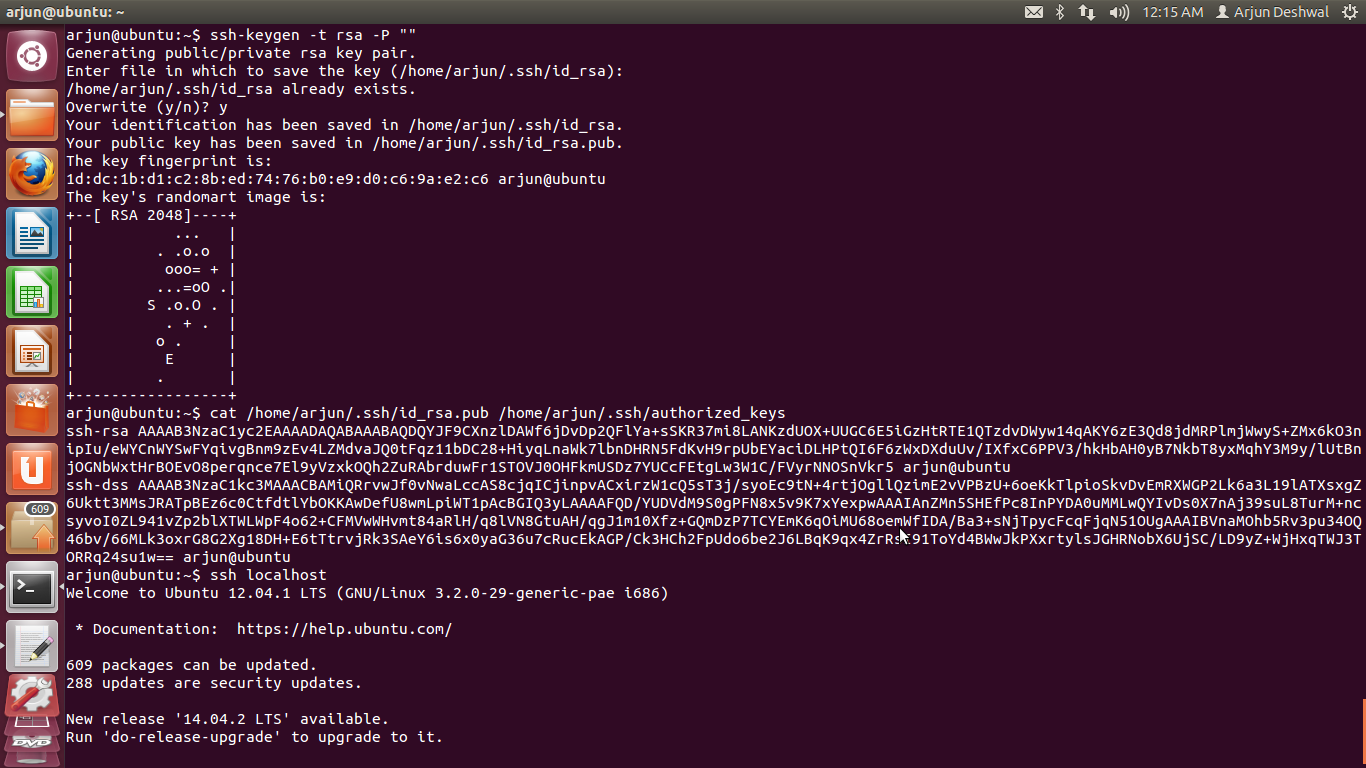
command:-ssh-keygen -t rsa -P ""

3.3)configure password-less SSH

command:-cat $HOME/.ssh/id\_rsa.pub >> $HOME/.ssh/authorized\_keys

3.4)check by SSH to localhost

command :-ssh localhost



1. Download Hadoop

download Apache hadoop

1. Install Hadoop

5.1)The downloaded tar file is extracted.

command:- tar xzf hadoop-2.2.0.tar.gz

5.2) Goto the folder

command:- cd hadoop-2.2.0

1. Setup / Configure

6.1) Edit configuration file (hadoop-env.sh) and set JAVA\_HOME. **(Fig 1)**

6.1.1)open file (hadoop-env.sh) folder:-$HOME/hadoop.2.2.0/etc/hadoop/hadoop- env.sh

6.1.2)set path:-

command:-export JAVA\_HOME=/usr/lib/jvm/java-6-openjdk-i386/

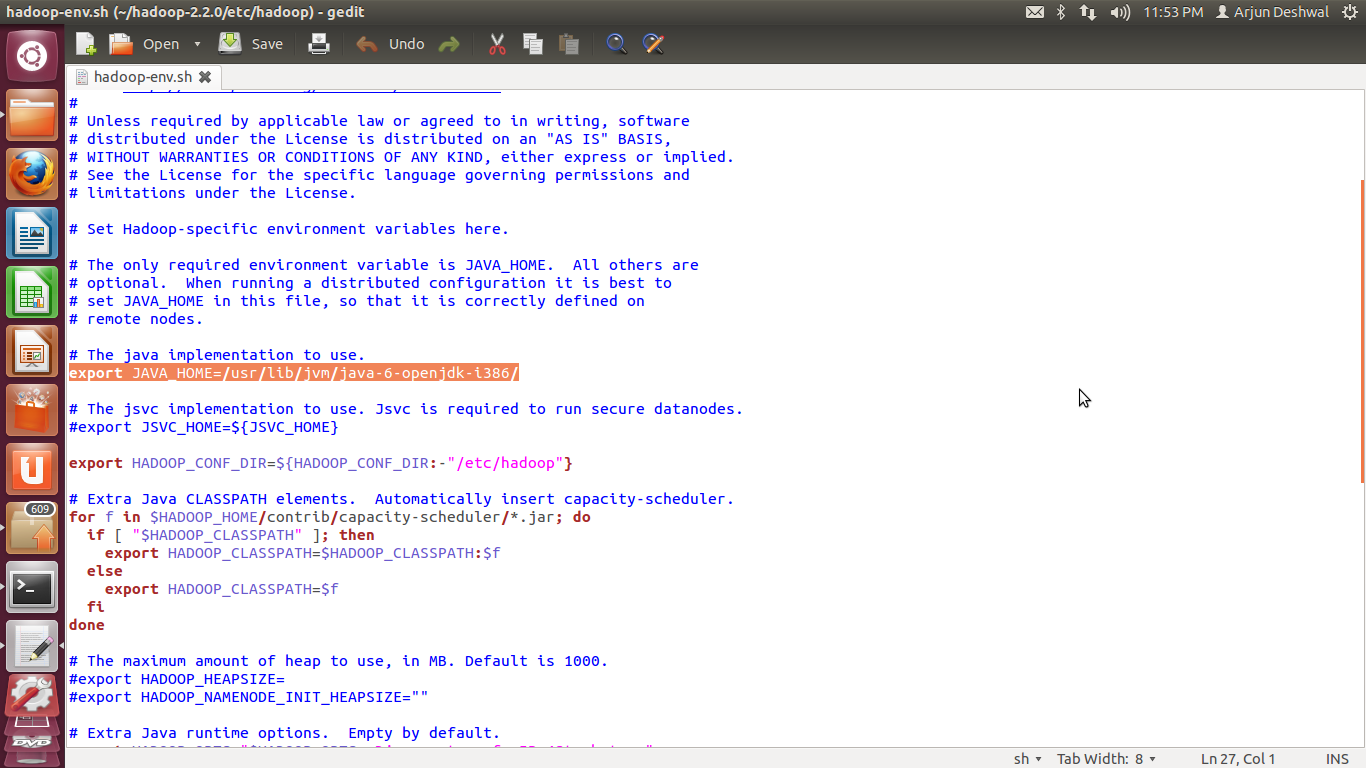


Fig 1

6.2) Edit configuration file (core-site.xml)(shown in **Fig 2):**

<configuration>

<property>

<name>fs.default.name</name>

<value>hdfs://localhost:9000</value>

</property>

</configuration>

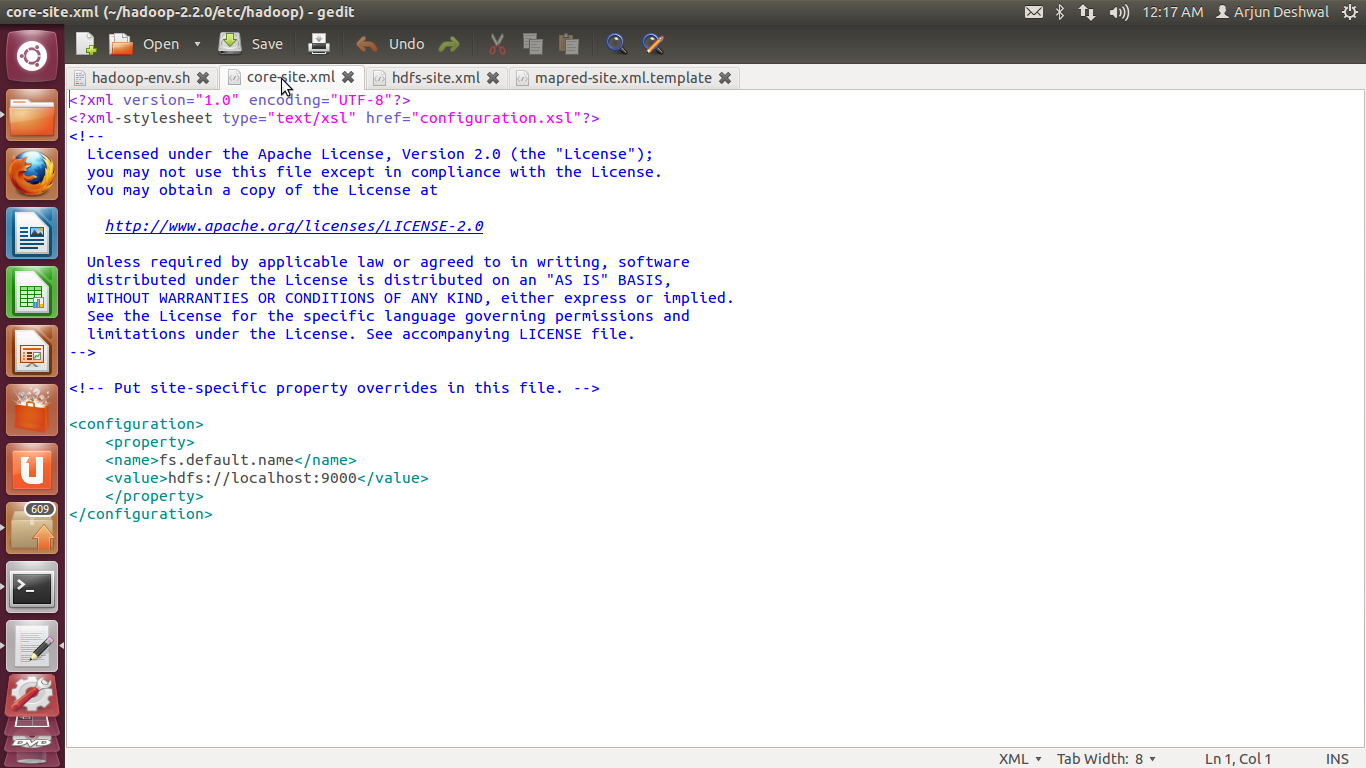


Fig 2

6.3) Edit configuration file (hdfs-site.xml) (shown in **Fig 3**)

<configuration>

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

</configuration>

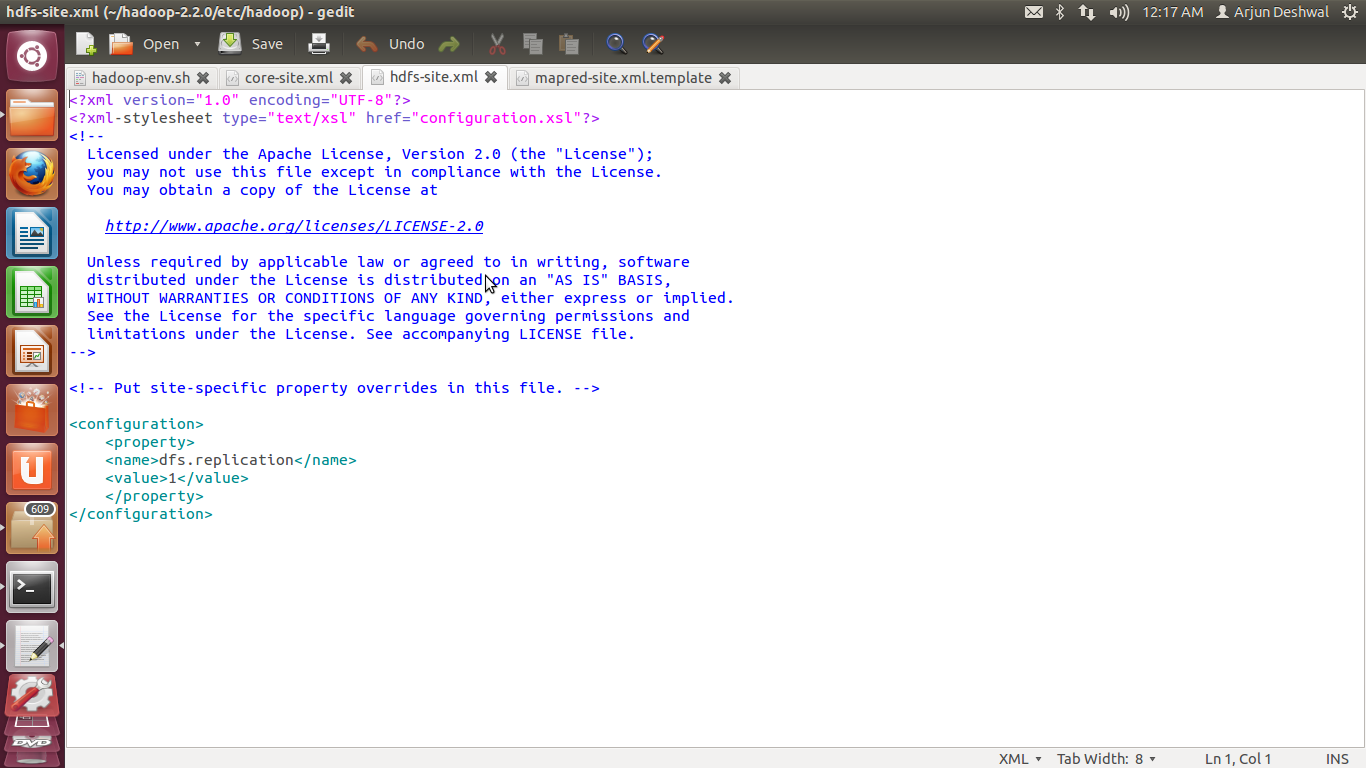


Fig 3

6.4) Edit configuration file (mapred-site.xml) (shown in **Fig 4**)

<configuration>

<property>

<name>mapred.job.tracker</name>

<value>localhost:9001</value>

</property>

</configuration>

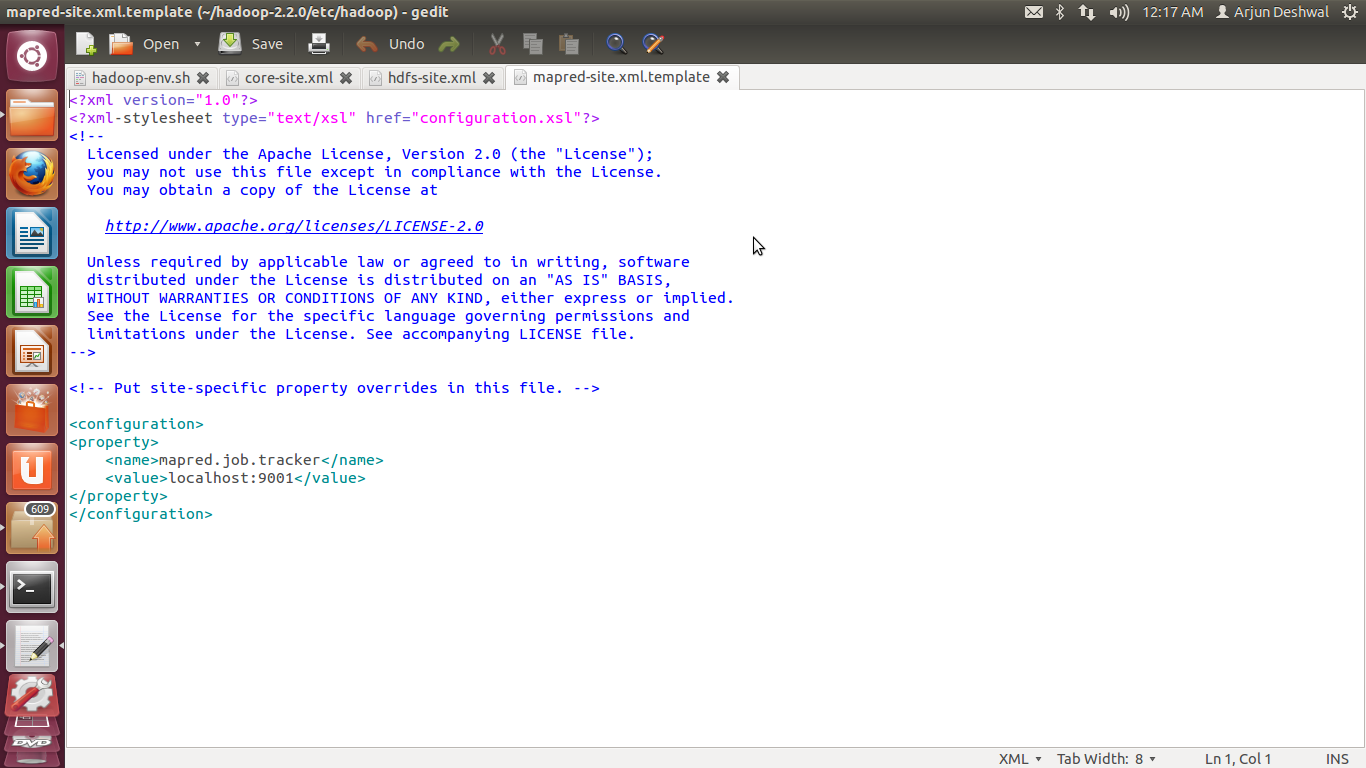
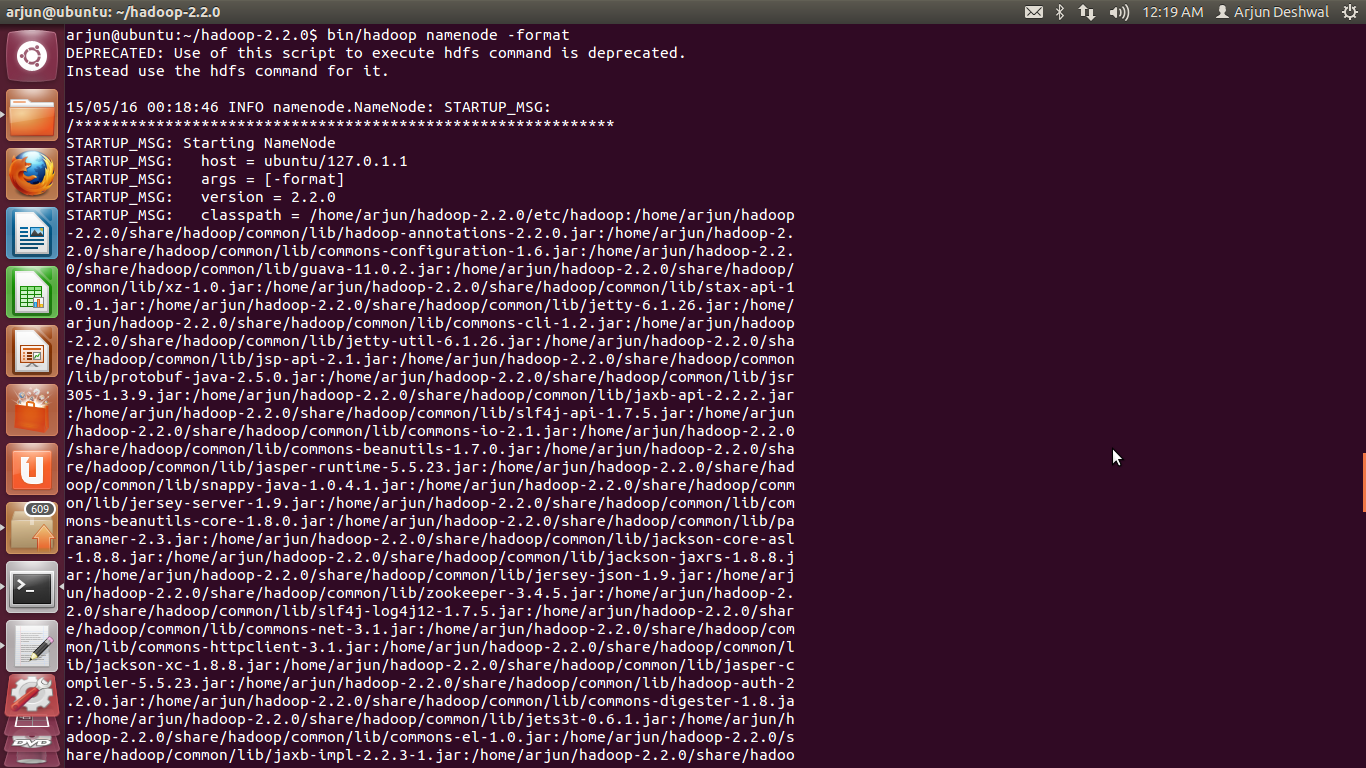


Fig 4

1. Start Hadoop cluster

7.1)Format Namenode(IMP :- this is done once when first time instaliing hadoop otherwise it will delete the data).

command:- bin/hadoop namenode -format



7.2)start Hadoop

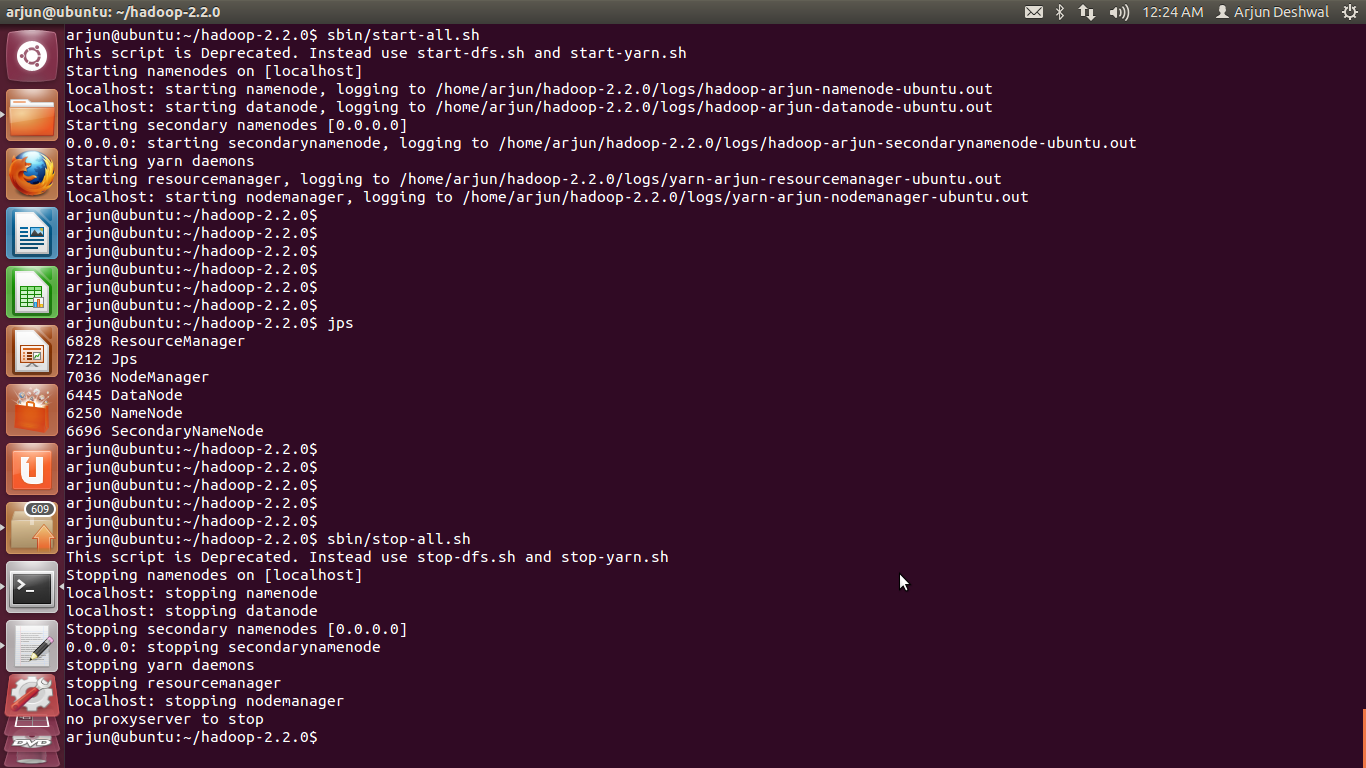
command:-sbin/start-all.sh

1. Check hadoop cluster running

command :-jps

1. Stop hadoop cluster

command:-sbin/stop-all.sh



**WEB CONSOLE VIEW**

Open Browser and type in the URL:-http://localhost:50070

