Set up and run hadoop 2.7.3

1. **prepare**

* The software should be prepared to install Hadoop 2.7.3 on windows 8.1 64bit
* Hadoop 2.7.3.tar.gz

(Link: <http://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-2.7.3/hadoop-2.7.3.tar.gz> )

* Java jdk 1.8.0.zip

(Link: <http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html> )

* Eclipse-standard-luna-R-win32-x86\_64.zip

(Link: <http://www.eclipse.org/downloads/packages/eclipse-standard-44/lunar> )

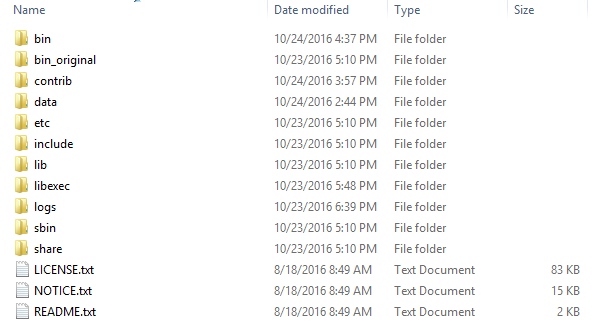
* Hadoop-eclipse-plugin-2.6.0.jar

(Link: <https://github.com/winghc/hadoop2x-eclipse-plugin/tree/master/release> )

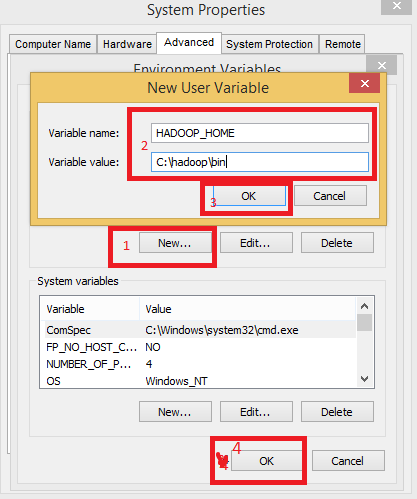
1. **Set up**

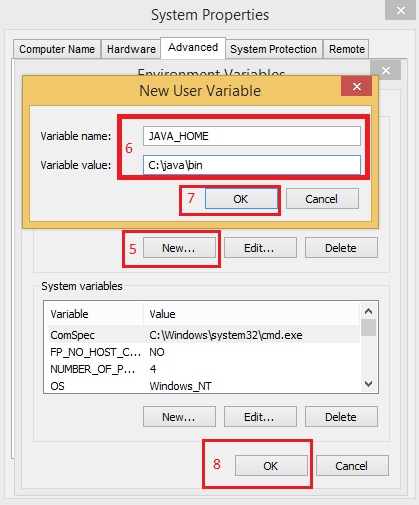
Step 1:

* Extract file Hadoop 2.7.3.tar.gz

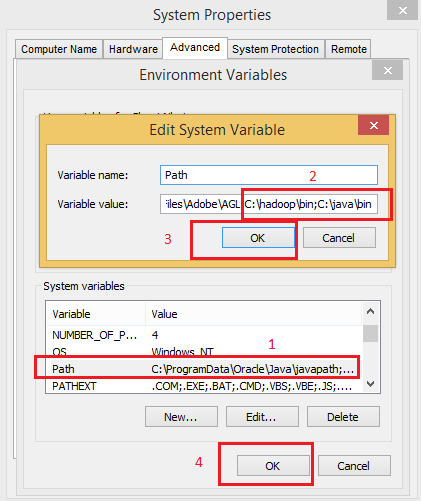


* Set the path **HADOOP\_HOME ,JAVA\_HOME Environment variable on windows ( Step 1,2,3…8 as 2 image below)**





* Next we set the directory path hadoop bin, java bin



Step 2: Configuration Hadoop

* Edit file hadoop-2.7.3/etc/hadoop/core-site.xml, paste the paragraph below and save.

<configuration>

<property>

       <name>fs.defaultFS</name>

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

   </property>

</configuration>

* Edit file hadoop-2.7.3/etc/hadoop/mapred-site.xml, paste the paragraph below and save.

<configuration>

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

</configuration>

* Edit file hadoop-2.7.3/etc/hadoop/hdfs-site.xml, paste the paragraph below and save.

<configuration>

<property>

       <name>dfs.replication</name>

       <value>1</value>

   </property>

   <property>

       <name>dfs.namenode.name.dir</name>

       <value>C:\hadoop-2.7.3\data\namenode</value>

   </property>

   <property>

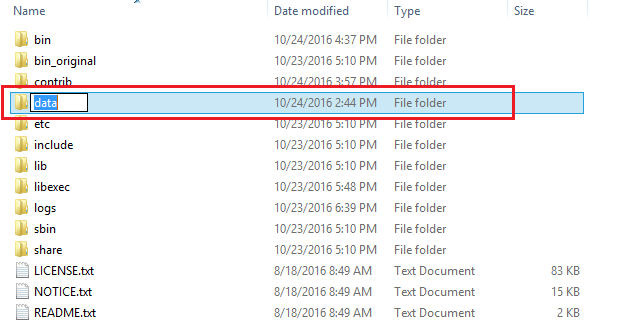
       <name>dfs.datanode.data.dir</name>

     <value>C:\hadoop-2.7.3\data\datanode</value>

   </property>

</configuration>

* Create folder data on Hadoop



* Edit file hadoop-2.7.3/etc/hadoop/yarn-site.xml, paste the paragraph below and save.

<configuration>

   <property>

       <name>yarn.nodemanager.aux-services</name>

       <value>mapreduce\_shuffle</value>

   </property>

   <property>

      <name>yarn.nodemanager.auxservices.mapreduce.shuffle.class

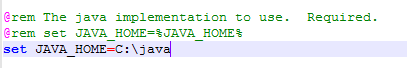
</name>

<value>org.apache.hadoop.mapred.ShuffleHandler</value>

   </property>

</configuration>

* Edit file hadoop/etc/hadoop/hadoop-env.cmd by closing the command line JAVA\_HOME=%JAVA\_HOME% instead of set JAVA\_HOME=C:java (On C:java this is path to file jdk.18.0)

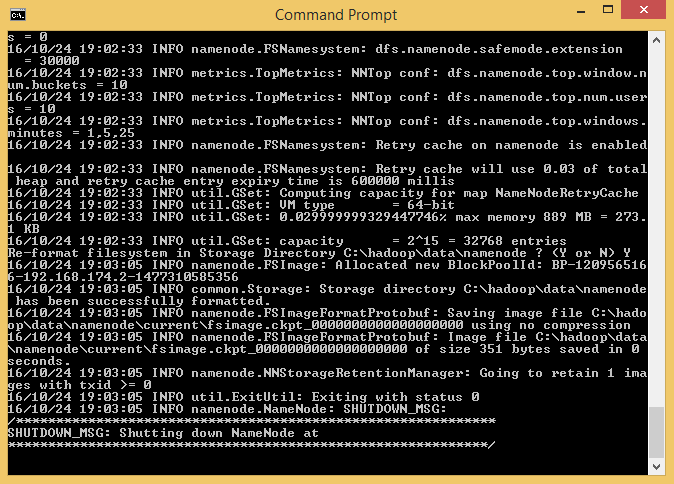


Step 3:

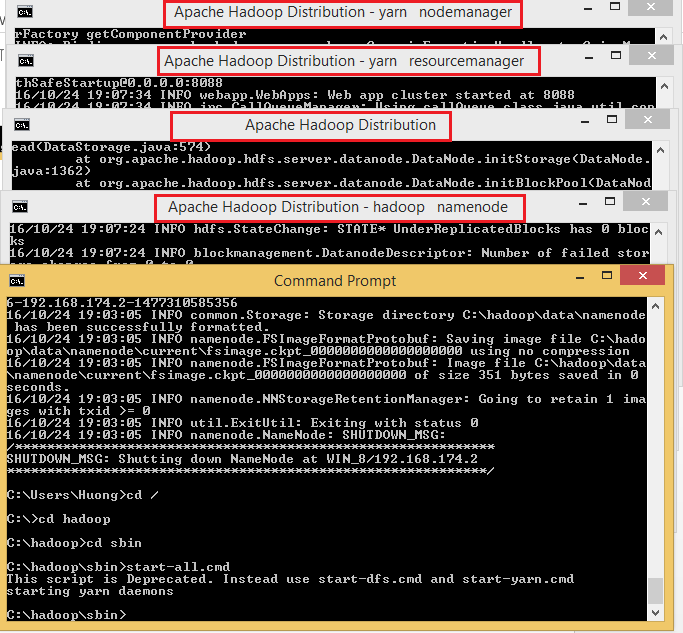
* Next: Dowload file sardetushar\_gitrepo\_download link: <https://github.com/sardetushar/hadooponwindows/archive/master.zip>

Next, Delete file bin on C:\hadoop\bin , replaced by file bin on file just download.

* Open cmd and typing command “hdfs namenode –format” . You will see

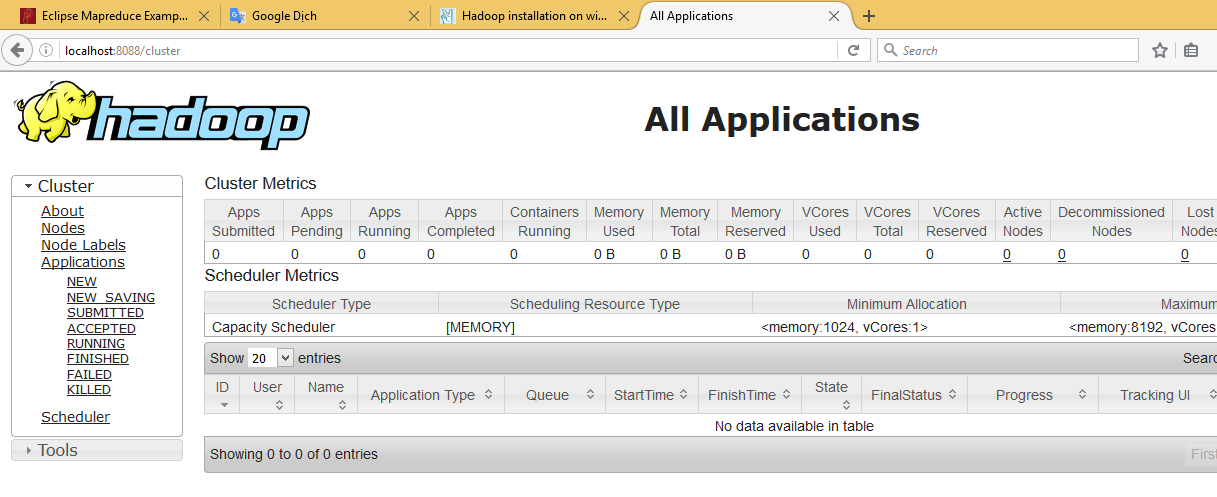


* Use cmd to sbin and typing “start-all.cmd” to start apache.
* Finish typing jps to check out
* Namenode
* Datanode
* YARN resourcemanager
* YARN nodemanager

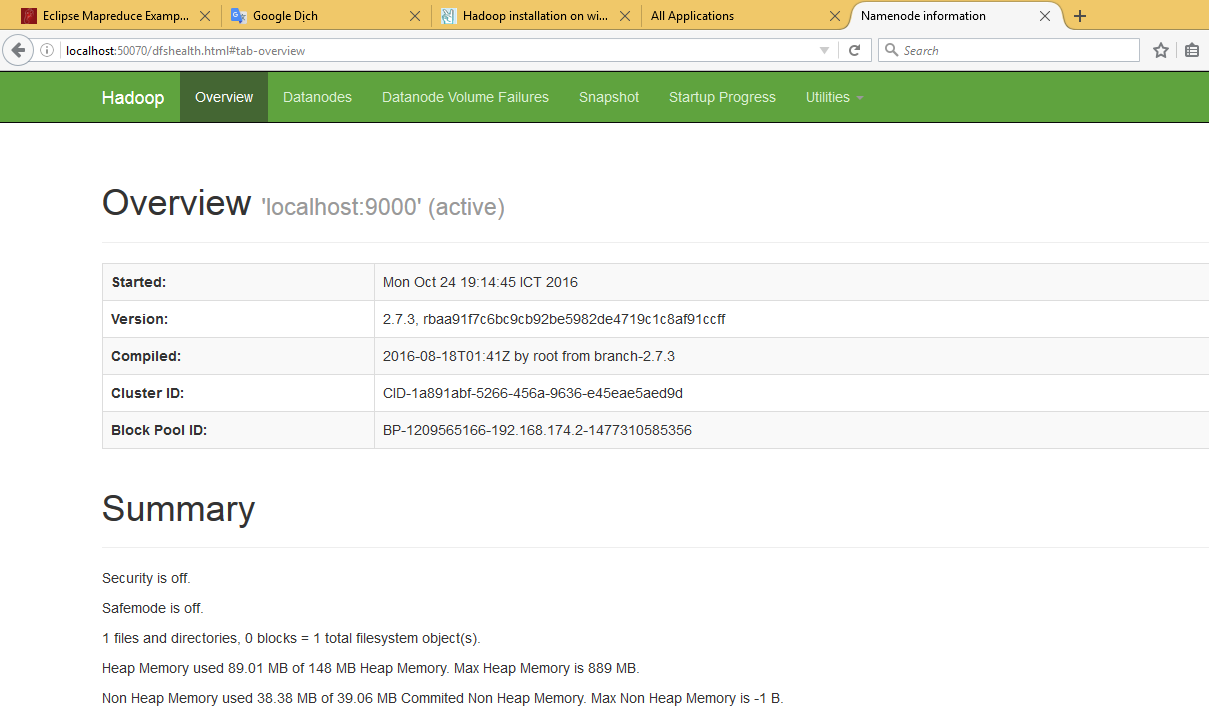


Step 4:

* Open: [http://localhost:8088](http://localhost:8088%20)



* Open: [http://localhost:50070](http://localhost:50070/)



So, Hadoop installed.

Note:

If in yarn resource manager and in yarn nodemanager there happens an error: Couldn't find a package.json file in "C:[\\hadoop-2.7.6\\sbin](../../../..//sbin)" (and there is no json file in that location).

Solution: There is a conflict with yarn module package dependency management. You can uninstall Yarn/Yarnpkg (delete Yarn folder, for ex. In C:\Yarn)