**Step-by-Step Hadoop Installation in Windows**

Pre-Requisites:

1. Install Java 1.8 (preferably 64 Bit)
2. Configure JAVA\_HOME using windows environment variables
3. Configure path for JAVA/bin directory

Steps for Hadoop Installation:

1. Download Hadoop-2.7.3 binary version, (Note: Binary version is pre-compiled) from apache website

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

Note: You might get tar file once you extract, so you need to extract the tar file until you get the extracted Hadoop-2.7.3 directory

1. Extract the Hadoop-2.7.3.tar.gz file into your Hadoop installation path, in my case C:/
2. Download following files
   1. Winutils.exe
   2. Hadoop.dll
   3. Hdfs.dll

From below website

<https://github.com/steveloughran/winutils/tree/master/hadoop-2.7.1/bin>

1. Copy the above files winutils.exe, hadoop.dll, hdfs.dll into the Hadoop bin directory; in my case,

C:/hadoop-2.7.3/bin/

1. Configure Hadoop home in your environment variables as below
   1. Open any windows folder
   2. Right Click on “This PC”
   3. Select properties
   4. Click Advanced System Settings
   5. Click on Environment Variables
   6. Under “user variables” section
   7. Click New
   8. Add variable name as “HADOOP\_HOME”
   9. Add variable value as your Hadoop installation folder, in my case “C:/hadoop-2.7.3”
   10. Click OK
   11. Again click on New
   12. Add another variable name as “HADOOP\_CONF\_DIR”
   13. Add variable value as etc/Hadoop directory in my case it is; “C:/hadoop-2.7.3/etc/hadoop”
   14. Click OK
   15. Under “System variables” section
   16. Select “Path” variable and click on Edit
   17. Click on New and add new entry for Hadoop/bin directory in my case; “C:/hadoop-2.7.3/bin”
   18. Now click ok and exit from the variables section
2. Go into Hadoop installation path under directory “C:/hadoop-2.7.3\etc\hadoop” and edit following files
3. Edit core-site.xml, add the required properties as below;

Note: Before editing core-site.xml file; create the directory; in my case “C:/hadoop-2.7.3/base”

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://localhost:8020/</value>

</property>

<property>

<name>hadoop.tmp.dir</name>

<value>C:/hadoop-2.7.3/base</value>

</property>

<property>

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

<value>hdfs://localhost:50071</value>

</property>

</configuration>

1. Edit hdfs-site.xml file and add the below properties;

Note: Make sure that you create all the directories configured in below properties, for e.g., C:/hadoop-2.7.3/base/dfs/data, C:/hadoop-2.7.3/base/dfs/name, C:/hadoop-2.7.3/base/dfs/secondaryname

<property>

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

<value>file:///C:/hadoop-2.7.3/base/dfs/data</value>

</property>

<property>

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

<value>file:///C:/hadoop-2.7.3/base/dfs/name</value>

</property>

<property>

<name>dfs.namenode.checkpoint.check.period</name>

<value>3600</value>

</property>

<property>

<name>dfs.namenode.checkpoint.txns</name>

<value>40000</value>

</property>

<property>

<name>fs.checkpoint.dir</name>

<value>file:///C:/hadoop-2.7.3/base/dfs/secondaryname</value>

</property>

<property>

<name>dfs.namenode.num.checkpoints.retained</name>

<value>1</value>

</property>

<property>

<name>mapreduce.jobtracker.restart.recover</name>

<value>true</value>

</property>

1. Need to do the name node format, for this
   1. -> Through CMD go to C:/Hadoop-2.7.3/bin/

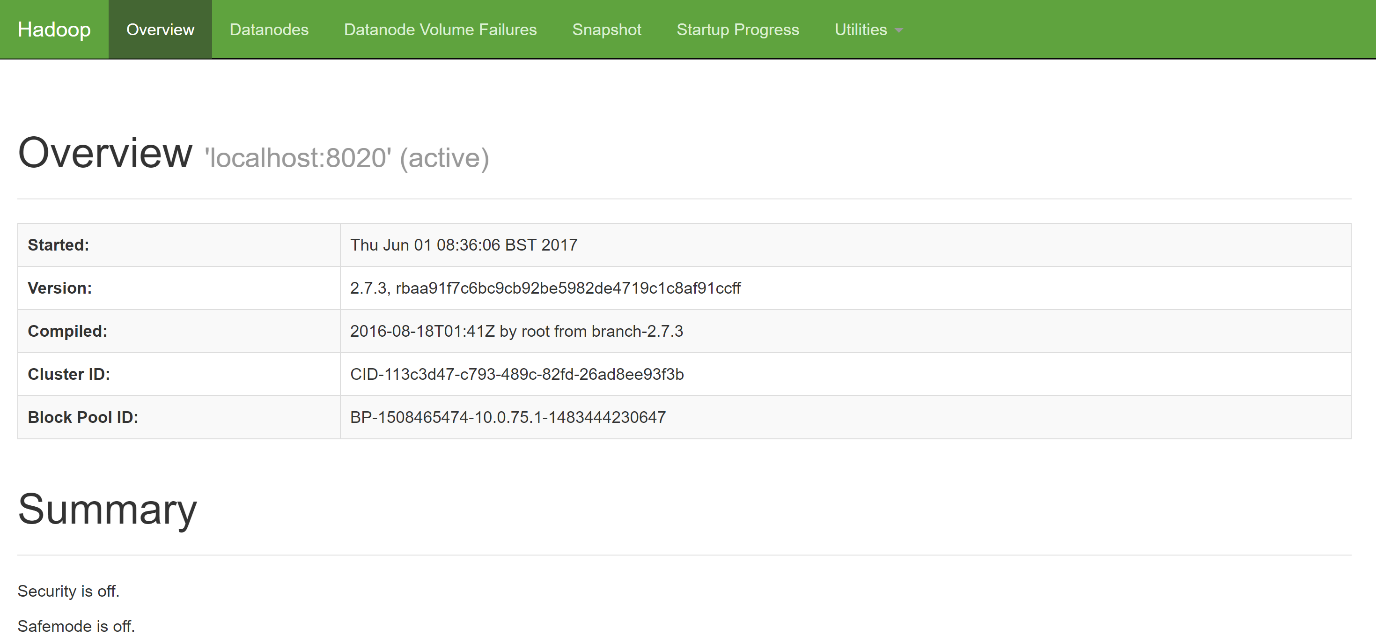
Run this command => hadoop namenode -format

Name node format should be completed successfully.

1. Open a new command prompt
2. Change your directory to Hadoop/sbin folder, in my case “C:/test/hadoop-2.7.3/sbin”
3. Run command “start-all”
4. Open the browser (like chrome, I.E, Firefox)
5. Access below URL

<http://localhost:50070/>

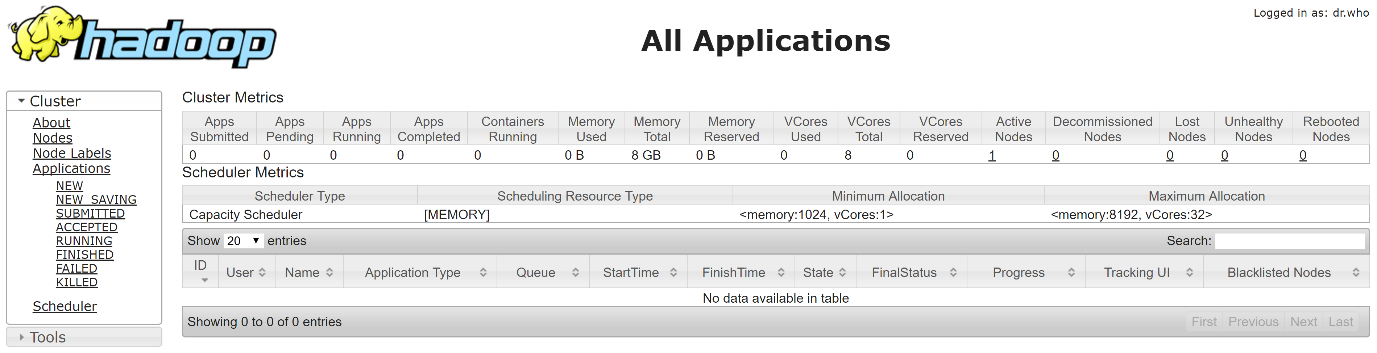
You must see the HDFS UI, check below screen shot



1. Access YARN-Resource manager URL as below

<http://localhost:8088/cluster>

You must see below UI,



1. Your Hadoop installation is Successful, Enjoy!