**Assignment:** Introduction to Hadoop

**Apache Hadoop Website**

<http://hadoop.apache.org/>

**Cloudera Quick Start VM -installation**

<http://www.cloudera.com/content/www/en-us/documentation/enterprise/latest/topics/cloudera_quickstart_vm.html>

**Hadoop Commands**

<http://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-common/FileSystemShell.html>

**Hue:** Hue is the open source web interface to Hadoop

**Login Credentials**: User name and password is cloudera

**Hello World program**

* **Create a simple file and put in HDFS**

|  |  |
| --- | --- |
| **Type This** | **It Means** |
| touch testfile.txt | create an empty file name testfile.txt |
| echo "Hello world in HDFS" >> testfile.txt | display the line of text that is being inserted into the newly created testfile |
| cat testfile.txt | read and print out the file |
| hdfs dfs -put testfile.txt | put the file into HDFS |

**Home Assignment**

Now that you have uploaded and installed the Hadoop VM and have experienced creating a file you should go a step further and make some changes to the existing file. You should use the Hue environment and copy the "Hello world in HDFS" line ten times. Save the newly edited file under the same name and place.

Take a look how large is the file and report back the number of KB this newly updated file size.