HDFS exercise:

HDFS Java API In this exercise,

you will have a chance to develop code that copies files to and from HDFS. You will utilize Hadoop’s Java API. Approx Time: 45 minutes Perform

1. Develop Java code that will copy files from HDFS to local filesystem. Your requirements are:

a. Copy files from HDFS directory '/training/exercises/filesystem/'

b. Copy files to local directory '/home/hadoop/Training/play\_area/exercises/filesystem/e1'

c. Only copy text files. In other words, only files that end with .txt

2. Develop Java code that will copy file from local filesystem to HDFS. Your requirements are:

a. Copy text files from local directory '/home/hadoop/Training/play\_area/exercises/filesystem/e2/' (files that end with .txt)

b. Copy files to HDFS directory '/training/playArea/filesystem/e2'

c. Report status to the screen by printing dots: '...'

* Understand daemon processes (Namenode, Secondary Namenode, Datanode)
* Commands to stop and start HDFS daemons
* Copying data back and forth to HDFS
* Understand parameter files and data files
* Restore and recovery of Namenode
* Important parameters and their defaults (dfs.blocksize, dfs.replication)
* Namenode Web UI