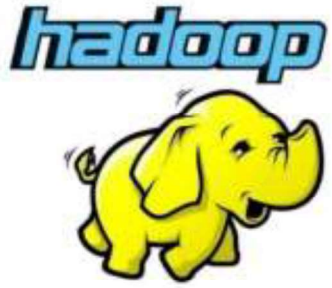
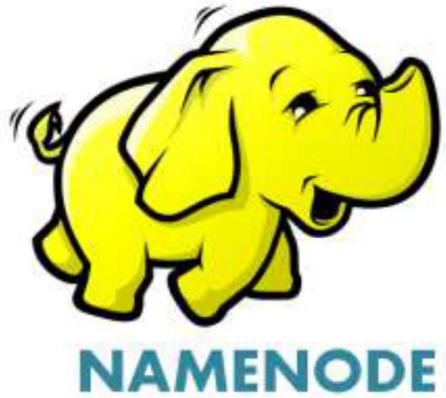
रा.इ.सू.प्रौ.सं
NIELIT



National Institute of Electronics and Information Technology

Big Data and Hadoop

Module 3-Starting with Hadoop



Hadoop Startup Scripts

The ${HADOOP_HOME}/sbin directory contains some scripts used to launch Hadoop DFS and Hadoop Map/Reduce daemons. These are:

Script Name	Purpose
start-dfs.sh	Starts the Hadoop DFS daemons, the namenode and datanodes. Use this before start-mapred.sh
stop-dfs.sh	Stops the Hadoop DFS daemons.
start-mapred.sh	Starts the Hadoop Map/Reduce daemons, the jobtracker and tasktrackers.
stop-mapred.sh	Stops the Hadoop Map/Reduce daemons.
start-all.sh	Starts all Hadoop daemons, the namenode, datanodes, the jobtracker and tasktrackers. Deprecated; use start-dfs.sh then start-mapred.sh
stop-all.sh	Stops all Hadoop daemons. Deprecated; use stop-mapred.sh then stop-dfs.sh

Start the Hadoop Cluster

First, start the NameNode and DataNode with the following command:

```
start-dfs.sh
```

You should get the following output:

```
Starting namenodes on [0.0.0.0]
Starting datanodes
Starting secondary namenodes [ubuntu2004]
```

Next, start the YARN resource and nodemanagers by running the following command:

```
start-yarn.sh
```

You should get the following output:

```
Starting resourcemanager
Starting nodemanagers
```

Start the Hadoop Cluster

To check the Hadoop services are up and running use the following command:

```
jps
```

You should get the following output:

```
5047 NameNode
5850 Jps
5326 SecondaryNameNode
5151 DataNode
```

HDFS Basic Commands

SNo	Command & Description
1	**-ls <path>** Lists the contents of the directory specified by path, showing the names, permissions, owner, size and modification date for each entry.
2	**-lsr <path>** Behaves like -ls, but recursively displays entries in all subdirectories of path.
3	**-du <path>** Shows disk usage, in bytes, for all the files which match path; filenames are reported with the full HDFS protocol prefix.
4	**-dus <path>** Like -du, but prints a summary of disk usage of all files/directories in the path.
5	**-mv <src><dest>** Moves the file or directory indicated by src to dest, within HDFS.
6	**-cp <src> <dest>** Copies the file or directory identified by src to dest, within HDFS.
7	**-rm <path>** Removes the file or empty directory identified by path.
8	**-rmr <path>** Removes the file or directory identified by path. Recursively deletes any child entries (i.e., files or subdirectories of path).
9	**-put <localSrc> <dest>** Copies the file or directory from the local file system identified by localSrc to dest within the DFS.
10	**-copyFromLocal <localSrc> <dest>** Identical to -put