Lab 11

# Intro:

Hadoop is an open-source framework that is used to store large amounts of data on any commodity hardware. This lab requires installation and configuration of hadoop and then running a MR job.

# Installation & Configuration:

Firstly install java on Ubuntu.

Also install openssh.

Then edit the hadoop-env.sh file located in conf folder for defining JAVA\_HOME.

# Execution of MR Job:

Run the following commands:

$bin/haoop namenode –format

$ bin/start-all.sh

Copy input files using

$ bin.haoop fs –put conf input

Run examples using :

$ bin/haoop jar example\*.jar grep input output ‘dfs[a-z.0+’

# Results:

The result output can be checked using the following command:

$ bin/hadoop fs –cat output/\*