Running the WeatherTemperature example

Summary

This guide describes how to compile and run the WeatherTemperature program, a Java program implementing the map and reduce functions used to analyse the logs from the National Climatic Data Center (NCDC). The logs present the temperature information of a given year. The WeatherTemperature program analyse the data do present the highest recorded global temperature for each year in the dataset.

Compiling the WeatherTemperature program

The WeatherTemperature program resides inside the WeatherTemperature folder. The folder is composed of the following files:

- src/version1/**MaxTemperatureMapper.java** - Contains the 'generical' map function implementation.
- src/version1/**MaxTemperatureReducer.java** - Contains the 'generical' reduce function implementation.

- src/version2/**MaxTemperatureMapper.java** - Contains the specific map function implementation (verify if the temperature is missing, if so, it applies a standard value).
- src/version2/**MaxTemperatureDriver.java** - Contains the execution main function of the program. This file define 'who' is the Mapper and the Reducer.

- **weather-temperature.jar** - The executable MaxTemperature file.

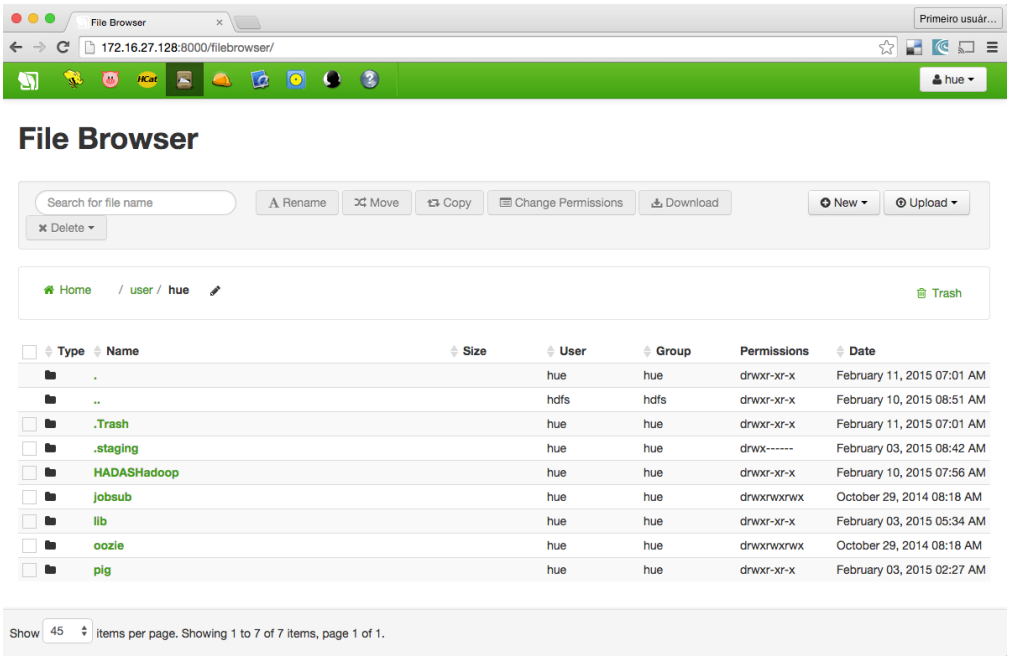
- inputs/* - The input files the be executed by the program.

- src/lib/*.jar - The hadoop library necessary to execute the program

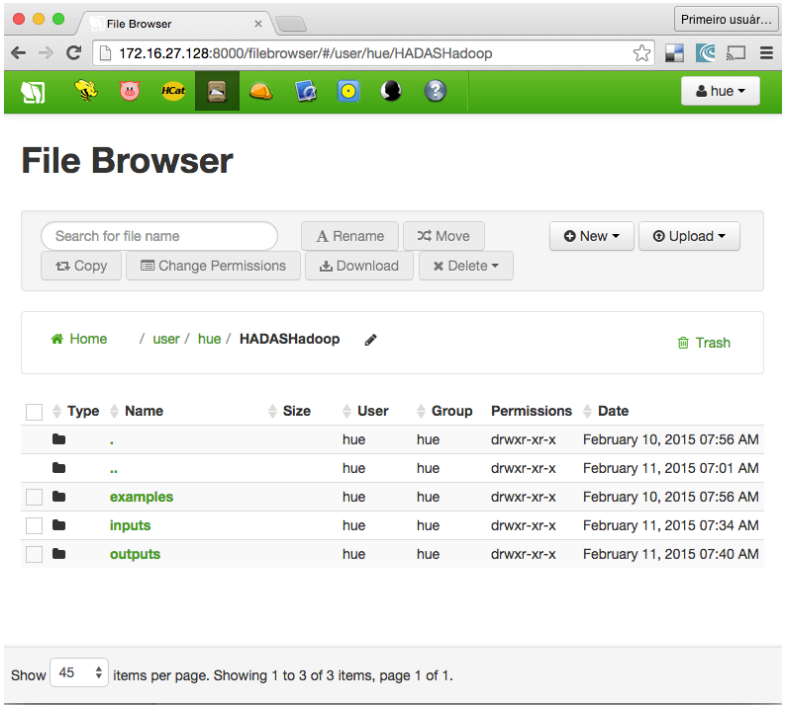
Inside order to compile the WeatherTemperature program, you can execute it in your prefered Java Tool (like Eclipse), or execute it in the command line. The program is already compiled in the Java version 1.7 and compressed in the **weather-temperature.jar** file, ready to be executed.

Running the WeatherTemperature program in Hadoop - HortonWorks

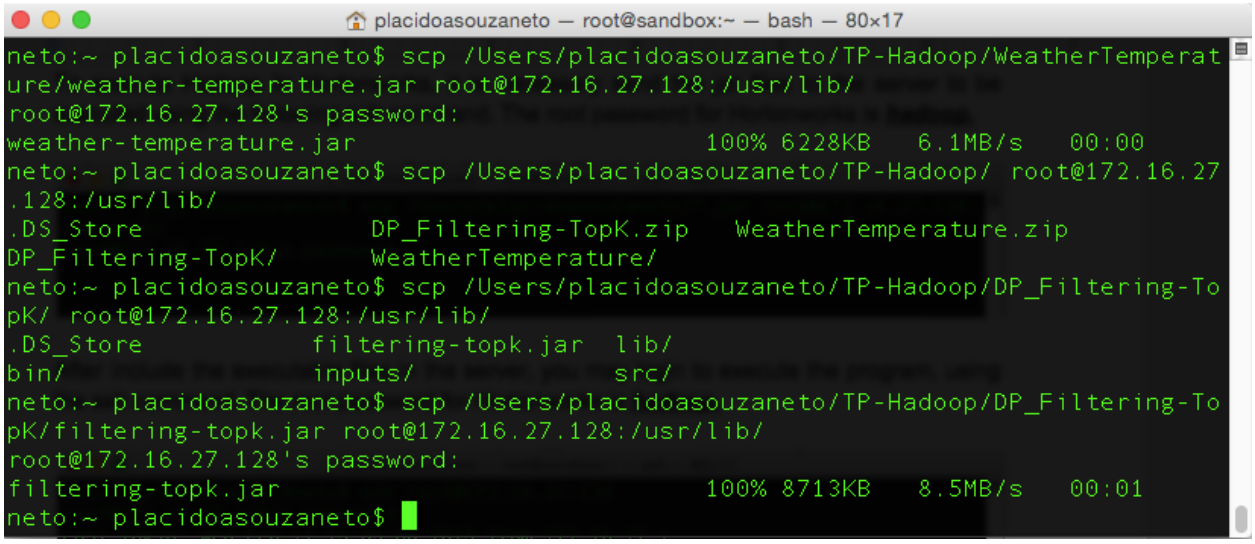
First you may access the HDFS filebrowser (http://172.16.27.128:8000/filebrowser/) in create a folder for the examples, for example: **HADASHadoop**.



In this folder you may also create 3 new folders: **inputs** and **outputs:**

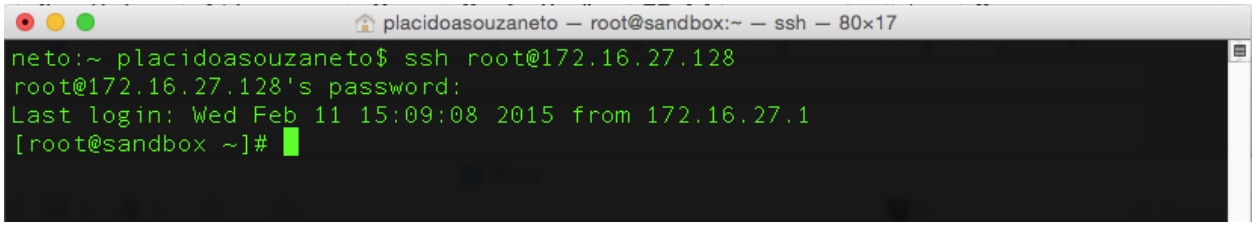


In the **inputs** folder you may include all the input files. The outputs files will be generated after the hadoop execution.



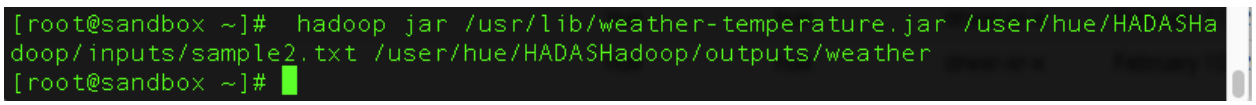
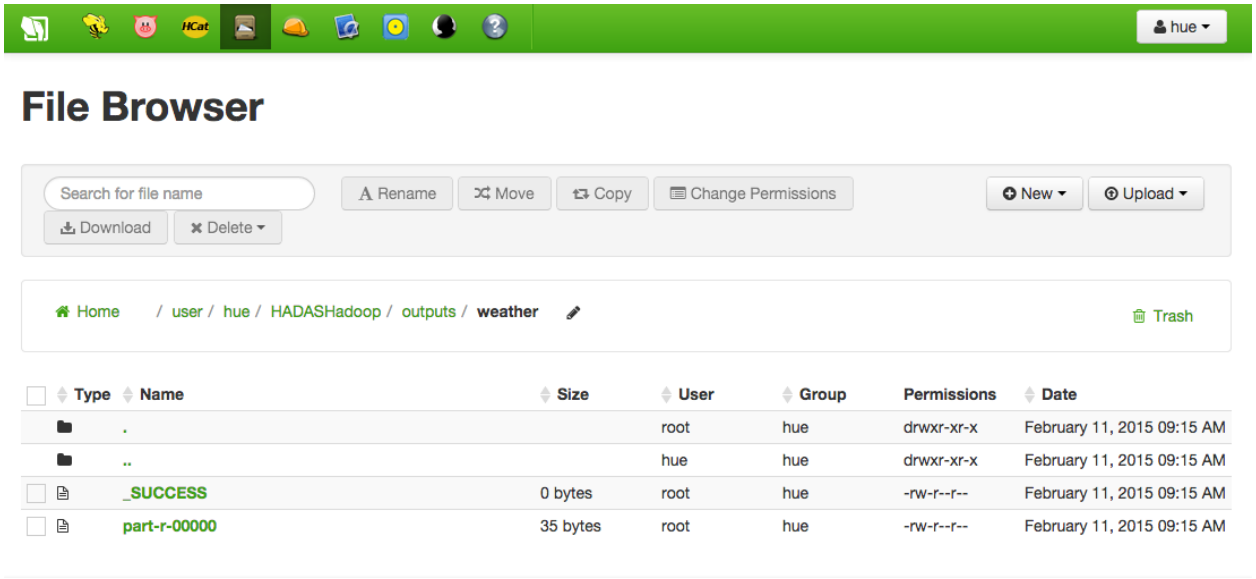
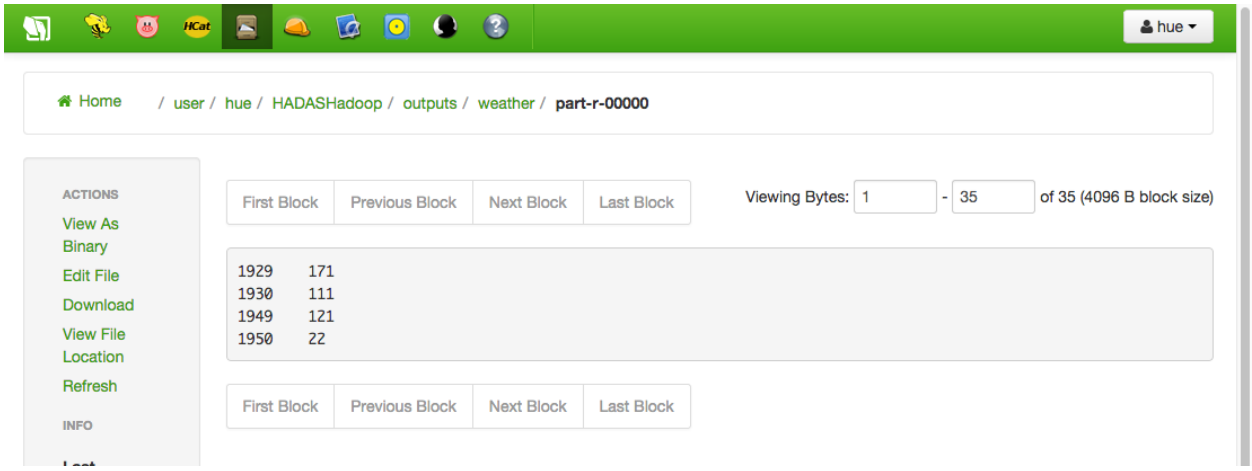
To execute Hadoop In Hortonworks, it is necessary send the jar files to the server to be executed using the following ssh command. The root password for Hortonworks is ***hadoop.***

After include the executable files in the server, you may login to execute the program, using the ssh command. The root password for Hortonworks is ***hadoop.***



Once Logged in Hadoop Server, you must execute the hadoop command:

hadoop jar *<ssh 'local' jar file> <hdfs input path> <hdfs output path>*



After that, the results can be reached in the HDFS file browser.