

Assignment 4

CSC241-Object Oriented Programing

Maximum Marks 20

Assigned Date: 26-12-2022

Due Date: 30-12-2022

Description

Design and implement a simple Java program to find specific strings in a large text. The program should be composed of the following modules:

1. The main module - reads a large text from all files in parts (e.g., 1000 lines in each part) and sends each part (as a string) to a matcher. After all, the word count is completed, it calls the aggregator to combine and print the results
2. The word counter - gets a text string as input and counts word frequency in the given string. The result is a map of a word to its location(s) in the text.
3. The aggregator - aggregates the results from all the word counts and prints the results. Please use the text files provided in the data folder for this task.
4. After completing the previous three tasks, use two computers to make the work count efficient. For this purpose create a socket connection between two PCs and calculate the running time of the entire execution. In addition, an equal number of files be distributed among both computers.
5. The extended aggregator combines the results from both computers and shows the aggregated results.
6. Display the running time without using threads, when using threads, and finally, using multiple machines. Write a conclusion of your findings.

Example of one line from the program output based on the input above:

Timothy --> 100

There should be several concurrent word counters (i.e., each counter should run in a separate thread). The results should be printed (in no particular order) after all text pieces have been processed. Please provide the main method that executes a sample run. You can make use of any written or web resource you find, incl. 3rd party libraries). The code should work and be as clean and efficient as you think necessary

Submission:

Submit Java source code zip file. The name of the folder should be your registration number.