

Homework 5: Parallel Word Search

In this homework, you will write a parallel program that performs a word search. More specifically, you will receive an input file containing a large grid of letters (lowercase). Your job will be to write a parallel program that finds all instances of a particular word in the grid.

All of the steps below must be performed on the head node of the memviz cluster (memviz.memphis.edu). That is, you must **ssh** into the head node to do the steps.

Step 1. Check out the project

You must first checkout the homework files. To do so:

1. SSH into the cluster's head node with this command:
ssh YOUR_ID@memviz.memphis.edu
2. Check out the **homework5** files with this command:
svn co https://utopia.cs.memphis.edu/course/comp4040-2013fall/uuids/YOUR_ID/homework5/

A folder **homework5** should have been downloaded. Recall that you can see a listing of the contents of the current directory using the command **ls**. Also you can change directories with the command **cd**. For example, you might do “**cd homework5**” to move into the **homework5** directory.

Step 2. Make and run the “main” program

To help get you started, you are given a program (**main.c**) that reads in a letter-grid file (**/home/sdf/homework5-input/letter_grid_small.txt**) and prints out the second row of letters (with each letter printed on a separate line). You may build your word-search program by adding to/modifying **main.c**. Also, you may practice searching the small letter grid file before switching to the big one (**/home/sdf/homework5-input/letter_grid_big.txt**), which has 10,000 rows and 10,000 columns. You can **cat** these files to view their contents.

To try out the **main** program, **cd** into the **homework5/** folder. Then, run these commands:

1. **make** – to compile the program in **main.c**
2. **make run** – to run the program

Recall that you can use the **nano** command to edit text files, such as **main.c**.

Step 3. Extend the “main” program

Your main task this project is to change the **main.c** program as follows:

- Use the **letter_grid_big.txt** file as input.
- Search the file for all occurrences of the word “pl” (short for *programming languages*), but search only the *horizontal rows* of the file.

- Print a count of how many times the word is found.
- Divide up the work among processes such that each process handles a different set of rows.
- One process (**rank 0**) should be different from the others in that it is responsible only for summing all the values returned by the workers and printing the total at the end.
- Your code should work correctly with any number of processes greater than 2.

Step 4. Submit your work

To submit, simply commit your files to the SVN repository. To do so, use the command:

svn -m “*My log message*” commit

Note that if you add new files, the above will not commit them. You must first use the **svn add** to add files to version control before you can commit them.

Extra Credit

To earn 3 points of extra credit (+15% added to the homework), make your program also search vertical columns (reading top to bottom). As before, you should divide the work up among your processes in a sensible way.