Exercises to work out and turn in:

Grading Guidelines (See Appendix):

In general, a right answer will get full credit when:

1. It is right (worth 25%)
2. It is right **AND** neatly presented making it easy and pleasant to read. (worth an **extra** 15%)
3. There is an **obvious and clear link[[1]](#footnote-1)** between 1) the information provided in the exercise and in class and 2) the final answer. A clear link is built by properly writing, justifying, and documenting an answer (worth an **extra** 60%).
4. Calculation mistakes will be minimally penalized (2 to 5% of full credit) while errors on units will be more heavily penalized.

**Late Submission** : as specified in the syllabus. Day counting starts one minute after the deadline.

**Check Your Submission:**  after submitting, download your submission to check whether it is the right version and it is complete.

You are welcome/encouraged to discuss exercises with other groups or the instructor. But, ultimately, **personal** writing is expected.

* USE THIS FILE AS THE STARTING DOCUMENT YOU WILL TURN IN. **KEEP IN THE QUESTIONS** AND **INSERT** YOUR ANSWERS **RIGHT AFTER THE QUESTIONS**.
* IF USING HAND WRITING (STRONGLY DISCOURAGED), REWRITE THE QUESTIONS.
* FAILING TO FOLLOW TURN IN DIRECTIONS /GUIDELINES WILL COST A 30% PENALTY.

Objectives of this assignment:

* to work on a Unix based system
* to become familiar with the notion of a *process*
* to assess/evaluate the *fairness* and *reliability* of the CPU scheduler on Tux machines

What you need to do:

* Examine and understand the pseudocode of a provided C source program *lab3*. (You do not need to dig into the C code)
* Run experiments of *lab3*
* Collect and process data about the experiments
* Build plots
* Interpret and discuss the plots
* Conclude about the fairness and reliability of the CPU scheduler.

**Important:**

* *One submission per team.*
* *Writing and presentation of your report are considered to grade your hands-on lab. Your conclusions* ***must be supported*** *by the data/measurements you collect. Your conclusions must be correct.*
* ***Questions about this lab must be posted on Piazza if you need timely answers****.*
* ***Work ahead. Do not wait until the last minute.***

Exercise (100 points)

This hands-on lab has multiple parts: 1) Compile, execute, skim, and understand the pseudocode of a provided C *program*, 2) run experiments, 4) collect data, and 5) analyze and report your results.

The instructor designed and implemented in C the provided C program lab3.c.

**The Provided C Program**

The help you understand the C program, the instructor provides here its pseudocode:

**Inputs**:

**np**: number of **concurrent** processes to execute

**nm**: number of minutes each process must run

**Output**:

Displays when each process starts executing its task

Displays the index achieved by the task executed by each process. The more a process "computes" (uses the CPU), the higher is c.

Collects (writes to a file) (done by each child process)

Displays the amount of time the process "ran" on the CPU (done by each child process)

Process the data collected by each process. (Done by Parent process)

Displays mean, standard deviation (std) and percentage std/mean of (Done by Parent process)

**Code (Parent)**:

for i=1 to ***np***

create (fork) a process Pi

for i=1 to ***np***

"green light" Pi to start its task

Compute/displays mean and standard deviation/mean of collected by child processes.

**Code of Process Pi (Child Process):**

wait for green light

Display start time

Execute A computing task for ***nm*** minutes (increases )

Displays the amount of time the process competed for the CPU (should be ***nm*** minutes)

Collects (writes to a file "file.dat")

Do not hesitate to ask questions or seek help on Piazza. Start early to avoid getting stuck at the last time when your classmates and instructor are not available.

**What You Need To Do**:

1. Download the C program lab3.c from Canvas
2. Store lab3.c to your Auburn home directory lab3 that you created in M1: hands-on laboratory exercise. This directory's path is ~/nnnn/lab3 where ~ is your Auburn home directory path.
3. **Log on** any Tux machine
4. **Go** to your lab3 directory at ~/nnnn/lab3 (See M1: Hands-on Lab..)
5. **Type** **ls** to insure the file *lab3.c* is in the current directory (i.e., ~/nnnn/lab3)
6. **Type** the command **cc -o lab3 lab3.c -lm** (This will compile lab3.c and produce executable lab3)
7. **Insure** that the compilation is successful. Check whether the executable lab3 is in the current directory (use ls -al). If successful, complete the following actions and answer the questions:
8. **Test** the program lab3 for running three child processes for 2 minutes
9. To test lab3, **type** the command ./lab3 3 2
10. **Observe** the outputs produced by lab3 execution
11. After you become familiar with lab3, you will need to **check** whether lab3 is a good measurement tool meeting some requirements described below.
12. After checking the requirements, you will run lab3, collect data, plot your results, and discuss your results

Showing the lab ran 3 processes for 2 minutes as the initial test before continuing on to the other requirements

A computer screen with white text

Description automatically generated

**Requirements (for lab3 to be a good measurement tool)**

1) **Requirement I:** If we run one process multiple times for the same number of **nm** minutes, the number should be the same each time within

2) **Requirement 2:** If we run **np** processes multiple times for ***k\* nm*** minutes, the number displayed should be times larger than the number displayed for running **nm** minutes within . For example, if I run 20 processes for minutes, the number displayed should be times larger than the number displayed for running 20 processes for minutes within .

3) **Requirement 3:** If we run **100**processes multiple times for ***2*** minutes, the number displayed should be times larger than the number displayed for running **1000** processes for **2**minutes within . For this requirement, your code will be tested with these values **100** and **1000**.

**CPU Scheduling Evaluation:**

The objectives are:

a) to check whether lab3 program meets the requirements

b) to evaluate the **fairness** and the **reliability.**

We describe below the experiments you must conduct to achieve the objectives.

a) Checking whether your task meets the requirements

**Requirement 1** (10 points):

- Run lab3 five times during 2 minutes with np = 1 (one process). For each run of lab3, collect and report here the five values you collected for .

Report here ....

Values for c during each run:

1. Process 0 c= 141200

2. Process 0 c= 141094

3. Process 0 c= 141189

4. Process 0 c= 141046

5. Process 0 c= 140814

A screen shot of a computer

Description automatically generated

This shows all 5 runs and their values along with the date command.

A screenshot of a computer program

Description automatically generated

Are the five values within of the mean of the 5 values of c?

Answer here .... Justify your answer

Mean = (141200 + 141094 + 141189 + 141046 + 140814) / 5 = 141068.6

Each value of c falls within the range calculated based on the mean ±10%, indicating that the values are within the specified tolerance level. Therefore, the requirement that the values of c should be within ±10% of the mean is met for these runs.

**Requirement 1I** (10 points):

**First**, Run lab3 once during 2 minutes with np = 20 (20 processes).

Collect the mean (for all 20 processes) .

Report it post a screenshot[[2]](#footnote-2) documenting the mean and percentage standard deviation/mean of your experiment

A black screen with white text

Description automatically generated

**dat**

Report it..post a screenshot documenting the average and percentage standard deviation/mean of your experiment

A screenshot of a computer

Description automatically generated

Screenshot showing the 20 processes running for 6 minutes. A number of the processes were filled with 0 within the first run, as with the second only processes 19 and 20 were filled with 0. I included this with the second screenshot just to show the processes running and their values.

Is within?

Answer here, justify.

For the two-minute run (first run) the mean for c sub 2 was equal to 54806.40.

For the six-minute run (second run) the mean for c sub 6 was equal to 270783.20.

Now to check if c sub 6 is equivalent to 3 times c sub 2:

3 x c2 = 3 x 54806.40 = 164419.20

Then the difference between c sub 6 and 3 x c sub 2:

|c6 – 3 x c2| = |270783.20 – 164419.20| = 106364.00

From this we deduct that this difference is greater than 10 percent of 3 x c sub 2:

10% x (3 x c2) = 0.1 x 164419.20 = 16441.92

Since 106364.00 > 16441.92, c sub 6 is not within 10 percent of 3 x c sub 2.

The calculated difference exceeds the allowed 10% variation indicating that the condition is not met. Therefore, the experiment does not satisfy this requirement.

**Requirement 1I1** (10 points):

**First**, Run lab3 once during 2 minutes with np = 100 (100 processes).

Collect the mean (for all 100 processes)

report here.. post a screenshot documenting the average and percentage standard deviation your experiment

A screenshot of a computer

Description automatically generated

I cannot show a date command or extension of the username jce0039 with the TUX number because of the number of processes (100) showing data on the command prompt. This does show the values for the mean and STDV percentage as requested.

**Second**, Run lab3 once during 2 minutes with np = 1,000 (1000 processes).

Collect the mean (for all 1,000 processes)

Report: post a screenshot documenting the average and percentage standard deviation/mean of your experiment

A screenshot of a computer

Description automatically generated

Same as before with to many processes to also show a date command but I believe we have done what is required. Again this shows the mean and STDV for 1000 processes over 2 minutes.

Is within?

Answer here, justify.

No, c sub 100 is not approximately 10 x c sub 1000 within a ±10 percent window. The mean value for 100 processes is significantly higher than 10 times the mean value for 1000 processes. This discrepancy suggests that the performance of the program does not scale linearly with the number of processes. The difference in these values indicates that there might be inefficiencies or bottlenecks in the program when handling a larger number of processes.

a) (30 points) **Fairness**: If the CPU scheduler is **fair**, the measure **c** (number displayed by lab3) must be close for all child processes you run within the **same** experiment. We call an *experiment* the fact to run the program *lab3* **once**. We measure how the values c produced by each child process are close by computing the percentage of the standard deviation/mean: the lower is this value, the closer are the c values of the processes, and the fairer is the CPU scheduler. Ideally, all processes must use the CPU for the same duration and therefore reach the same value c.

In order to evaluate fairness, you must conduct one experiment during 2 minutes for each of the following values of **np** (number of processes): 2, 5, 10, 20, 40, 60, 80, 100, 200, 400, 500, 1000. For each experiment with value , compute the percentage of the standard deviation/mean of the measure c your task displays for each process. Plot the percentage of the standard deviation/mean of c versus the number of processes (put the np values 2, 5, 10, 20, 40, 60, 80, 100, 200, 400, 500, 1000 on the x-axis and the percentage of the standard deviation/mean of c on the y-axis) Discuss this plot and draw reasonable conclusions regarding fairness. In other words, is the CPU scheduler of the Tux machines fair based on the data you collected?

Based on the data we collected from running the requested experiments with different numbers of processes, it is obvious that the fairness of the CPU scheduler on the Tux machines varies significantly based on the number of processes involved. R2 = 0.9671, which means that approximately 96% of the variation can be explained by the number of processes (independent variable) in the linear regression model. This indicates a strong relationship between the variables highlighting that the CPU scheduler on the Tux machines is not fair as the number of processes increases. The fairness deteriorates significantly as the number of processes grows, leading to uneven distribution of CPU time among processes. This lack of fairness is evident from the high standard deviation/mean percentages observed in the experiments with larger numbers of processes. Anything less than 10 processes seems to keep fairness of the CPU scheduler. The CPU scheduler should ideally provide equal opportunities to all processes, ensuring that they receive similar CPU time, but the data indicates a substantial deviation from this ideal scenario with 1000 processes having a massive 2234.02 percent difference.

b) (20 points) **Reliable**: If the CPU scheduler is **reliable**, the average measure **c** (number displayed by lab3) for **np** processes must remain the same (within .) for successive experiments running for the same number of processes **np** each time. In other words, if you execute the same process on a machine, the execution time should not widely vary: the execution time should not be one time 2 minutes, then 2 hours, then 1 minute, then 30 minutes.....

In order to evaluate reliability, run 10 times lab3 with np=100 (100 process) during 2 minutes (./lab3 100 2). Report here the mean of c for each run.

Run 1 mean value of c: 1362.41

Run 2 mean value of c: 1027.65

Run 3 mean of c: 354.35

Run 4 mean value of c: 214.73

Run 5 mean value of c: 377.96

Run 6 mean value of c: 1175.09

Run 7 mean value of c: 2636.72

Run 8 mean value of c: 1614.64

Run 9 mean value of c: 1358.08

Run 10 mean value of c: 586.73

Plot here the mean versus the run # (put the values 1, 2, ....10 on the x-axis) and the mean on the y-axis. (no tables)

Discuss here this plot and draw reasonable conclusions regarding reliability. In other words, is the CPU scheduler of the Tux machines reliable based on the data you collected?

Based on the observed variability in the mean values of c across runs, it can be concluded that the CPU scheduler of the Tux machines is not entirely reliable. The lack of consistent execution times suggests that the CPU scheduler does not ensure uniform allocation of CPU resources for similar processes across different runs. R2 = 0.0583, which means that approximately 6% of the variation can be explained by the number of runs of conducted (independent variable) in the linear regression model. Highlighting that there is a weak relationship between the mean values of “C” over multiple runs. Runs 3 and 4 show significantly lower mean values compared to the other runs, indicating unusually short execution times. Runs 6 and 7 show considerably higher mean values, suggesting longer execution times. In summary, the data indicates that the CPU scheduler exhibits unreliability, as evidenced by the significant fluctuations in execution times observed during the 10 runs. Further analysis would be needed to find the underlying issues in the differences between not only the fairness but also the reliability of the CPU scheduler on the TUX machine.

**Do not hesitate to ask questions on Piazza if you have any doubt.**

**Getting Started**

1. compile the code I provided you by typing:

cc -o lab3 lab3.c -lm

1. Execute the code (evaluate 1 process during 2 minutes): ./lab3 1 2
2. Observe what is displayed
3. Compile and execute to check whether your task meets the requirements (see above)
4. Make sure to complete all required tasks.
5. Do not hesitate to ask questions on Piazza or during office hours.

**Common mistakes**

1) Finish the code at the last minute: you will not have enough time to collect data and write a good report

2) Finish the lab/report at the last minute without polishing may result in a low score. You must have time to reflect about the meaning of these experiments and the conclusions you may draw.

**What to turn in?**

One file:

**Electronic copy** of **this** file that includes your answers. I repeat: you must insert your answers and plots in **this** file. Do not delete anything from this file. This file with your plots and answers must be put posted **separately** on Canvas (not in a zipped folder). **A penalty of 30 points will be applied if these instructions are not followed.**

Good writing and presentation are expected.

**In case of doubt, do not hesitate to ask questions on Piazza.**

What you need to turn in:

* Electronic copy of this file (including your answers) (standalone). Submit the file as a Microsoft Word or PDF file.
* Recall that answers must be well written, documented, justified, and presented to get full credit.
* How this assignment will be graded:
* A right answer will get full credit when:
* It is right (worth 25%)
* It is right AND neatly presented making it easy and pleasant to read. (worth 15%)
* There is an obvious and clear link between 1) the information provided in the exercise and in class and 2) the final answer. A clear link is built by properly writing, justifying, and documenting an answer (worth 60%).
* Calculation mistakes will be minimally penalized (2 to 5% of full credit) while errors on units will be more heavily penalized.
* You are welcome/encouraged to discuss exercises with other students or the instructor. But, ultimately, personal writing is expected.

**Appendix**: Grading: What is an OBVIOUS and CLEAR LINK?

Here is an example to explain what an **obvious and clear link** is and how we grade your work.

Consider the following problem:

"(100 points) John travels from Auburn to Atlanta in his car at a speed of 50 mph. Leaving at 8am, at what time will John reach Atlanta".

Here are the answers of three students and their scores:

**Student 1** answers: "10am". Student 1 will get 25 points.

**Student 2**answers : "John will reach Atlanta at 10am". Student 2 will get 25+15 = 40 points

**Student 3** answers: "The time t to travel a distance d at speed v is equal to d/v = d/50mph. The problem does not provide the distance d from Auburn to Atlanta. Based on Google, the distance from Auburn to Atlanta is approximately 100 miles (**document is here**). Therefore, the time t = 100 miles/50mph = 2 hours. Since John left at 8am, he will then reach Atlanta at 8am + 2 hours = 10 am".

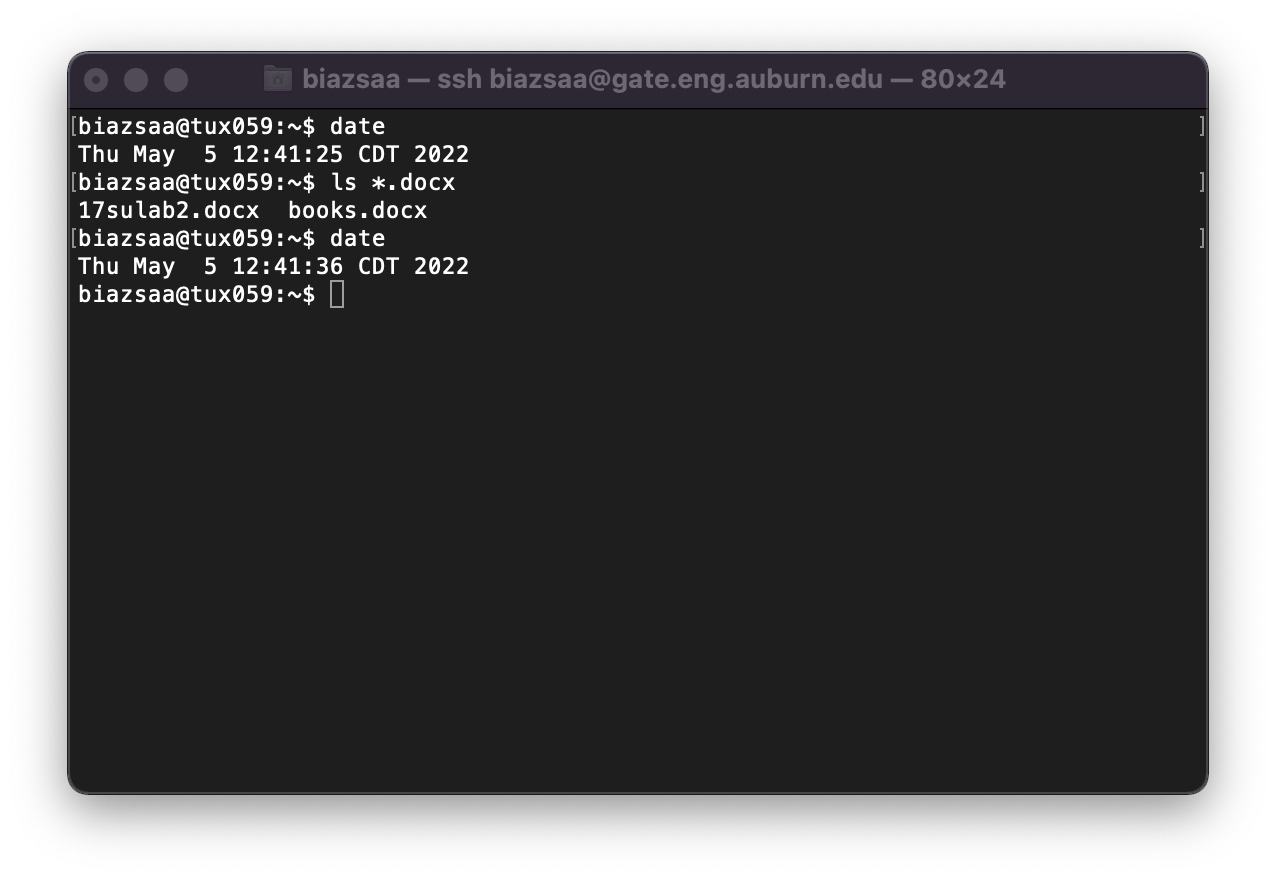
**Student 3** will get 25 + 15 + 60 = 100 points

Do you see the **direct** **link** going from the data provided in the question to the final answer, using general knowledge/formula and documents?.... Can you now solve the following problem and get 100 points?

"(100 points) Alice travels from Auburn to Atlanta in her car at a speed of 50 mph. Leaving at 8am, at what time will Alice reach Atlanta assuming that she had a flat tire that delayed her 30 minutes".

**Screenshot: Required Information**

**In order to save space, for this assignment and all *FUTURE* ones, clip out the screenshots to contain only the relevant information. *When applicable, ALL screenshots must show the date, the Tux machine you are using for the exercise and the Auburn username of one of the team mates*. Make sure that the screenshots are easily readable. Below is template screenshot:**

****

1. See on the appendix what an obvious and clear link is. [↑](#footnote-ref-1)
2. Review Appendix about the required information on ALL screenshots [↑](#footnote-ref-2)