CS 4550 OS Spring 2018 Scheduling Assignment

Due before class Tuesday Jan. 30, submit zipped or compressed files to ELearning.  Just the source code and make files, no object files, no executables.

You will write code to simulate a running operation system’s scheduling of processes.  Your code will be one process/thread only and the ‘processes’ are just data structures used in your simulation.

Make sure you read this and understand completely what you are doing before writing your design and code!

Rules for code:

A.     Must be C language supported by Visual Studio (VS) or CS department’s Ubuntu without installing any new compilers or libraries, as in standard C.  **Include the complete project zipped up with Make files and supporting files.**

B.     You must write your own priority queue code not use built in data structures so we can see how you chose to implement the algorithm.

C.      You must credit where you got the algorithms and code you use any part of.

D.  Use good modules (functions), your main loop should call functions and fit on one screen.

Do not try to do everything in one function but break it up into logical modules.

Write the hardest part first so you know what it needs.

Write the lowest level, test then write what calls it, test.  Do not try to write whole thing at once unless you really enjoy wasting your own time.

Here are the rules for all programs:

Rules:

If your code does not follow these rules you will have to rewrite it before you get a score.

        Test for an error return from ALL system calls, that includes **all**semaphore calls, fork, etc.

o   If error call perror() unless timer is up for timed wait. (not really an error)

        NO break or continue statements in any loop.

        No global variables, you declare, outside of a function.

        All loops must stop only when false, the Boolean expression in the ( ) after while, or for.

        All code must be properly formatted with indentation for all blocks { }

* Functions that are not void have one and only one return statement the last line in the function.
* Must have your name, class name, and assignment name in comments at top of file.
* Have comments that explain how things work if not completely obvious on first glance.  x—; // No decrement x comment  which is obvious

This simulation will assume one processor with one core.  It will assume all user processes with kernel taking no time to do its work.  How fake is that?

Time will be the **main loop one clock tick per iteration**.  One time through the loop one clock tick.

Priorities should be 1 to 20 with 1 lowest and 20 highest.

You must use the priority queue algorithm adjusting priorities each time you add a process to the ready queue.

You will also use the Round Robin algorithm so the process in the CPU is forced to leave when quantum time is up.

You will have 5 \* 20 processes, 5 for each priority from 0-20.  One that is CPU bound, one that is even CPU and I/O and 1 that is I/O bound, one between CPU and even and one between even and I/O bound. CPU bound must want more time than normally allowed in CPU before fast I/O spending most time in CPU.  I/O bound should stay in CPU much less time than normal CPU time and spend most time waiting for I/O.  Even should spend about the same amount of time in CPU as doing I/O.

There should be several data structures.

A priority queue of processes ready for the CPU.  Can be an Array constantly sorted by priority.

Each time click (one time around the main loop) you adjust priorities and move processes in this queue.

A list of processes waiting for I/O

A struct of OS parameters common to all processes

Max time in CPU before being bumped to ready queue if no I/O, quantum.

Max wait time user processes in ready queue (try different times, how small can your scheduling algorithm handle?)

Each process should keep track of **at least** the following information.

Name of process

Starting priority, reset current priority to this each time process moved into wait queue, does not change.

Current priority of the process which may change with aging.

A.     Time in CPU needed before/between I/O (set once)

B.     Time I/O takes (set once)

C.      Total time in machine not counting time in wait queue (set at startup counts down).  How much time process spends doing something before it exits.  Sum of total time in CPU and total time in I/O.

D.     Time in CPU currently.  Set to 0 when moved into CPU, when reaches A or quantum is up moves out of CPU.

E.      Time left waiting for current I/O

F.      Time process has been waiting in ready queue

G.     Total time in CPU

H.     Total time in I/O

I.       Total time in ready queue

J.       Smallest time in ready queue

K.      Longest time in ready queue

Other data

               You need to keep track of the current state of each process; you might have a field in each process or use which data structure it is in.  In CPU, I/O wait, Ready queue ….

You should have modules (functions() ) to

Swap process in CPU (how hard would it be to simulate this taking time as it does in real life?  Not required but think about changing your simulator for latency.  Context switch between kernel and user takes even longer in real system)

Add process to ready queue

Program.

Define at least 5 \* 20 processes with different starting times

Add one to the ready queue, run the loop at least once then add the next.

loop until all processes exit

Each iteration is one time unit.

Adjust time counters in each process

swap processes as necessary from CPU to either the I/O or ready queue

swap processes from I/O to ready queues when I/O time is up

Adjust priority of processes in ready queue so no processes waits too long and so higher priorities run first.

After the loop stops and all process have exited print the statistics for

               the OS setup for time limits

each process’s stats including total, min and max times

All I/O times should be > 2

Document your program:

Which scheduling algorithm**s** are you using?  Based on what information source?

How does the algorithm handle making sure all processes get a chance to run even with high priority CPU bound processes?

You can read about real simulations in your text book.  Your simulation is more to learn how schedulers work than evaluation of algorithms.

Explanation of the Round Robin Priority Queue with ageing algorithm, the required one.

Priority Scheduling

Priority scheduling is when which each job is assigned a priority and the job with the highest priority gets scheduled first.

Priority scheduling can suffer from a major problem known as indefinite blocking, or starvation, in which a low-priority task can wait forever because there are always some other jobs around that have higher priority.

If this problem is allowed to occur, then processes will either run eventually when the system load lightens (at say 2:00 a.m. ), or will eventually get lost when the system is shut down or crashes. (There are rumors of jobs that have been stuck for years. )

One common solution to this problem is aging, in which priorities of jobs increase the longer they wait. Under this scheme a low-priority job will eventually get its priority raised high enough that it gets run.

Round Robin Scheduling

CPU bursts are assigned with limits called time quantum.

When a process is given the CPU, a timer is set for whatever value has been set for a time quantum.

If the process finishes its burst before the time quantum timer expires, then it is swapped out of the CPU and either exits or is waiting for I/O

If the timer goes off first, then the process is swapped out of the CPU and moved to the ready queue.