Section 4: Scheduling and Synchronization

CS162

July 8, 2019

Contents

[Handwritten annotations at top of page: "A B C arrive @ sametime", "3 3 3", "TQ: 1"]

[Handwritten table:]

	1	2	3	4	5	6	7	8	9	A	B	C	Avg
FCFS	A	A	A	B	B	B	C	C	C	0	3	6	3
RR	A	B	C	A	B	C	A	B	C	4	5	6	5

1　Warmup

Which of the following are true about Round Robin Scheduling?

1. The average wait time is less that that of FCFS for the same workload. *[F]*

2. Is supported by `thread_tick` in Pintos. *[T]*

3. It requires pre-emption to maintain uniform quanta. *[T]*

4. If quanta is constantly updated to become the # of cpu ticks since boot, Round Robin becomes FIFO. *[T]*　*[FCFS]*

5. If all threads in the system have the same priority, Priority Schedulers **must** behave like round robin. *[F]*

6. Cache performance is likely to improve relative to FCFS. *[F]*

7. If no new threads are entering the system all threads will get a chance to run in the cpu every `QUANTA*SECONDS_PER_TICK*NUMTHREADS` seconds. (Assuming `QUANTA` is in ticks). *[solns]*

8. This is the default scheduler in Pintos *[T]*

9. It is the fairest scheduler

2　Vocabulary

- **Scheduler** - The process scheduler is a part of the operating system that decides which process runs at a certain point in time. It usually has the ability to pause a running process, move it to the back of the running queue and start a new process;

- **FIFO Scheduling** - First-In-First-Out (aka First-Come-First-Serve) scheduling runs jobs as they arrive. Turnaround time can degrade if short jobs get stuck behind long ones (convoy effect);

- **Round-Robin Scheduling** - Round-Robin scheduling runs each job in fixed-length time slices (quanta). The scheduler preempts a job that exceeds its quantum and moves on, cycling through the jobs. It avoids starvation and is good for short jobs, but context switching overhead can become important depending on quanta length;

- **Priority Scheduling** - Priority scheduling runs the highest priority job, based on some assigned priorities. Starvation of low-priority jobs and priority inversion (a higher priority task waiting for a lower priority one, usually for a lock) are issues. Priorities can be static or dynamic, and if dynamic can change based on heuristics or locking-related donations;

- **SRTF Scheduling** - Shortest Remaining Time First scheduling runs the job with the least remaining amount of computation time and is preemptive. It has the optimally shortest average turnaround time. In practice remaining computation time can't be predicted, so SRTF is often used as a post-facto benchmark for other algorithms; *[SJF]*

- **Multi-Level Feedback Queue Scheduling** - MLFQS uses multiple queues with priorities, dropping CPU-bound jobs that consume their entire quanta into lower-priority queues;

- **Condition Variable** - A synchronization variable that provides serialization (ensuring that events occur in a certain order). A condition variable is associated with:

  - a lock (a condition variable + its lock are known together as a **monitor**)

– some boolean condition (e.g. `hello < 1`)

– a queue of threads waiting for the condition to be true

In order to access any CV functions **OR** to change the truthfulness of the condition, a thread must/should hold the lock. Condition variables offer the following methods:

– **cv_wait(cv, lock)** - Atomically unlocks the lock, adds the current thread to **cv**'s thread queue, and puts this thread to sleep.

– **cv_notify(cv)** - Removes one thread from **cv**'s queue, and puts it in the ready state.

– **cv_broadcast(cv)** - Removes all threads from **cv**'s queue, and puts them all in the ready state.

When a **wait()**ing thread is notified and put back in the ready state, it also re-acquires the lock before the **wait()** function returns.

When a thread runs code that may potentially make the condition true, it should acquire the lock, modify the condition however it needs to, call notify() or broadcast() on the condition's CV, so waiting threads can be notified, and finally release the lock.

Why do we need a lock anyway? Well, consider a race condition where thread 1 evaluates the condition C as false, then thread 2 makes condition C true and calls **cv.notify**, then 1 calls **cv.wait** and goes to sleep. Thread 1 might never wake up, since it went to sleep too late.

• **Hoare Semantics** - (In terms of condition variable) Wake a blocked thread when the condition is true and transfer control of the CPU and ownership of the lock to that thread immediately. This is difficult to implement in practice and generally not used despite being conceptually easier to deal with.

```
if (condition){
    wait()
}
```

• **Mesa Semantics** - (In terms of condition variable) Wake a blocked thread when the condition is true, with no guarantee that the thread will execute immediately. The newly woken thread simply gets put on the ready queue and is subject to the same scheduling mechanisms as any other thread. The implication of this is that **you must check the condition with a while loop instead of an if statement because it is possible for the condition to change to false between the time the thread was unblocked and the time it takes over the CPU**.

```
while (condition)
    wait()
}
```

3 Problems

3.1 Locking Up the Floopies

In section 3, you may remember encountering race conditions inside of the Central Galactic Floopy Corporation's currency exchange server, which runs on top of pthreads. We said that we could make the transactions run correctly by making the balance increment/decrement atomic. The Central Galactic Floopy Corporation hires a consultant named Morty who suggests making the increment/decrement pair appear atomic by adding a lock to each account, and acquiring the locks when we run the transaction.

```
typedef struct account_t {
  pthread_mutex_t lock = PTHREAD_MUTEX_INITIALIZED;
  int balance;
  long uuid;
};

void transfer(account_t *donor, account_t *recipient, float amount) {
  // lock accounts so we can make the transfer safely
  pthread_mutex_lock(&donor->lock);
  pthread_mutex_lock(&recipient->lock);

  // check balances and make transfer if possible
  if (donor->balance < amount) {
    printf("Insufficient funds.\n");
  } else {
    donor->balance -= amount;
    recipient->balance += amount;
  }

  // unlock accounts
  pthread_mutex_unlock(&recipient->lock);
  pthread_mutex_unlock(&donor->lock);
}
```

to prevent deadlock, enforce a global ordering of lock preference

If we use the locking code given above, will the code run correctly? Has Morty introduced a new bug into our code? Can you give an example of where this code would fail?

Can you modify the code above to resolve this bug?

3.2 Scheduling

Consider the following single-threaded processes and their arrival times, CPU bursts, and priority

Process	Arrival Time	CPU Burst	Priority
A	1	5	1
B	3	3	3
C	5	2	2
D	4	4	4

Please note:

- Priority scheduler is preemptive.

- Newly arrived processes are scheduled last for RR. When the RR quanta expires, the currently running thread is added at the end of to the ready list before any newly arriving threads.

- Break ties via priority in Shortest Remaining Time First (SRTF).

- If a process arrives at time x, they are ready to run at the beginning of time x.

- Ignore context switching overhead.

- The quanta for RR is 1 unit of time.

- Total turnaround time is the time a process takes to complete after it arrives.

Fill in the following scheduling table and calculate the total turnaround time for each scheduling algorithm

Time	FIFO	RR	Quum	SRTF	Priority
1	A	A	A	A	
2	A	A	A	A	
3	A	A	AB	B	
4	A	B	BAD	B	
5	A 5	A	ADBC	B	
6	B	D	DBCA	C	
7	B	B		C	
8	B 6	C		A	
9	D	A 9		A	
10	D	D		A	
11	D	B		D	
12	D 9	C		D	
13	C	D		D	
14	C 10	D		D	
Total Turnaround Time	30	37		27	

3.3 Simple Priority Scheduler

We are going to implement a new scheduler in Pintos we will call it SPS. We will just split threads into two priorities "high" and "low". High priority threads should always be scheduled before low priority threads. Turns out we can do this without expensive list operations.

For this question make the following assumptions:

- Priority Scheduling is NOT implemented

5

- High priority threads will have priority 1

- Low priority threads will have priority 0

- The priorities are set correctly and will never be less than 0 or greater than 1

- The priority of the thread can be accessed in the field `int priority` in `struct thread`

- The scheduler treats the ready queue like a FIFO queue

- Dont worry about pre-emption.

Modify `thread_unblock` so SPS works correctly.
You are not allowed to use any non constant time list operations

```
void
thread_unblock (struct thread *t)
{
  enum intr_level old_level;
  ASSERT (is_thread (t));
  old_level = intr_disable ();
  ASSERT (t->status == THREAD_BLOCKED);


  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------
  list_push_back (&ready_list, &t->elem);


  -------------------------------------------------
  t->status = THREAD_READY;
  intr_set_level (old_level);
}
```

3.3.1 Fairness

In order for this scheduler to be "fair" briefly describe when you would make a thread high priority and when you would make a thread low priority.

6

3.3.2 Better than Priority Scheduler?

If we let the user set the priorities of this scheduler with `set_priority`, why might this scheduler be preferable to the normal pintos priority scheduler?

3.3.3 Tradeoff

How can we trade off between the coarse granularity of SPS and the super fine granularity of normal priority scheduling? (Assuming we still want this fast insert)

3.4 Totally Fair Scheduler

You design a new scheduler, you call it TFS. The idea is relatively simple, in the begining, we have three values `BIG_QUANTA`, `MIN_LATENCY` and `MIN_QUANTA`. We want to try and schedule all threads every `MIN_LATENCY` ticks, so they can get atleast a little work done, but we also want to make sure they run *atleast* `MIN_QUANTA` ticks. In addition to this we want to account for priorities. We want a threads priority to be inversely proportial to its `vruntime` or the amount of ticks its spent in the CPU in the last `BIG_QUANTA` ticks.

You may make the following assumptions in this problem:

- Priority scheduling in Pintos is functioning properly,
- Priority donation is not implemented.
- Alarm is not implemented.
- `thread_set_priority` is never called by the thread
- You may ignore the limited set of priorities enforced by pintos (priority values may span any `float` value)
- For simplicity assume floating point operations work in the kernel

3.4.1 Per thread quanta

How long will a particular thread run? (use the threads priority value)

7

3.4.2   struct thread

Below is the declaration of `struct thread`. What field(s) would we need to add to make TFS possible? You may not need all the blanks.

```
struct thread
  {
    /* Owned by thread.c. */
    tid_t tid;                          /* Thread identifier. */
    enum thread_status status;          /* Thread state. */
    char name[16];                      /* Name (for debugging purposes). */
    uint8_t *stack;                     /* Saved stack pointer. */
    float priority;                      /* Priority, as a float. */
    struct list_elem allelem;           /* List element for all threads list. */

    /* Shared between thread.c and synch.c. */
    struct list_elem elem;              /* List element. */
#ifdef USERPROG
    /* Owned by userprog/process.c. */
    uint32_t *pagedir;                  /* Page directory. */
#endif
    _____          /* What goes here? */

    _____          /* What goes here? */

    _____          /* What goes here? */

    /* Owned by thread.c. */
    unsigned magic;                     /* Detects stack overflow. */
  };
```

3.4.3   thread tick

What is needed for `thread_tick()` for TFS to work properly? You may not need all the blanks.

```
void
thread_tick (void)
{
  struct thread *t = thread_current ();

  /* Update statistics. */
  if (t == idle_thread)
    idle_ticks++;
#ifdef USERPROG
  else if (t->pagedir != NULL)
    user_ticks++;
#endif
  else
    kernel_ticks++;

  _____;

  _____;
```

```
  /* Enforce preemption. */
  if (++thread_ticks >= TIME_SLICE) {  /* TIME_SLICE may need to be replaced with something else */
    intr_yield_on_return ();

    _____;

    _____;

    _____;

    _____;

    _____;

    _____;

    _____;
  }
}
```

3.4.4   timer interrupt

What is needed for `timer_interrupt` for TFS to function properly.

```
static void
timer_interrupt (struct intr_frame *args UNUSED)
{
  ticks++;
  _____;

  _____;

  _____;

  _____;

  _____;

  _____;

  _____;

  _____;

  _____;

  _____;

  _____;

  _____;
```

```
  thread_tick ();
}
```

3.4.5   thread create

What is needed for `thread_create()` for TFS to work properly? You may not need all the blanks.

```
tid_t
thread_create (const char *name, int priority, thread_func *function, void *aux)
{
  /* Body of thread_create omitted for brevity */

 ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------

  ------------------------------------------------


  /* Add to run queue. */
  thread_unblock (t);
  if (priority > thread_get_priority ())
    thread_yield ();

  return tid;
}
```
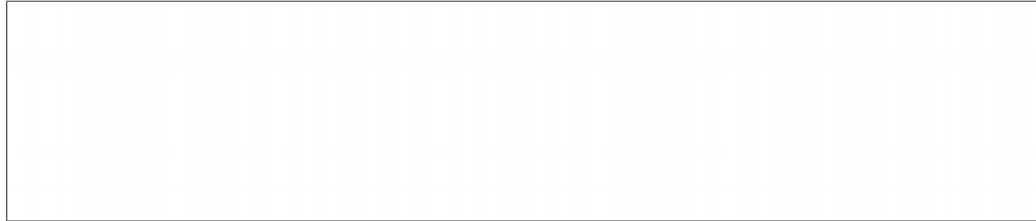
3.4.6 Analysis

Explain the high level behavior of this scheduler; what exactly is it trying to do? How is it different/similar from/to the multilevel feedback scheduler from the project?

3.5 test_and_set

In the following code, we use test_and_set to emulate locks.

```
int value = 0;
int hello = 0;

void print_hello() {
    while (test_and_set(&value));
    hello += 1;
    printf("Child thread: %d\n", hello);
    value = 0;
    pthread_exit(0);
}

void main() {
    pthread_t thread1;
    pthread_t thread2;
    pthread_create(&thread1, NULL, (void *) &print_hello, NULL);
    pthread_create(&thread2, NULL, (void *) &print_hello, NULL);
    while (test_and_set(&value));
    printf("Parent thread: %d\n", hello);
    value = 0;
}
```
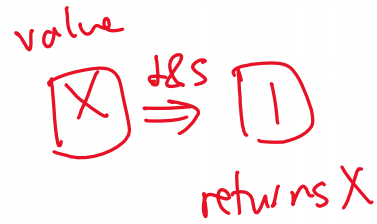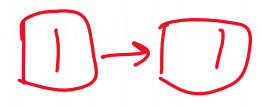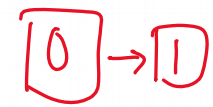
Assume the following sequence of events:
1. Main starts running and creates both threads and is then context switched right after
2. Thread2 is scheduled and run until after it increments hello and is context switched
3. Thread1 runs until it is context switched
4. The thread running main resumes and runs until it get context switched
5. Thread2 runs to completion
6. The thread running main runs to completion (but doesn't exit yet)
7. Thread1 runs to completion

Is this sequence of events possible? Why or why not?

At each step where `test_and_set(&value)` is called, what value(s) does it return?

Given this sequence of events, what will C print?

Is this implementation better than using locks? Explain your rationale.

3.6 Hello World

This code compiles (given a sprinkling of `#include`s etc.) but doesn't work properly. Why?

```
pthread_mutex_t lock;
pthread_cond_t cv;
int hello = 0;

void* print_hello(void* arg) {
    hello += 1;
    printf("First line (hello=%d)\n", hello);
    pthread_cond_signal(&cv);
    pthread_exit(0);
}

int main() {
    pthread_t thread;
    pthread_create(&thread, NULL, print_hello, NULL);
    while (hello < 1) {
        pthread_cond_wait(&cv, &lock);
```

(Handwritten annotations:)

lock_acquire(&lock)

lock_release(&lock)

← forgot to initialize lock

← did not acquire lock before calling wait

lock_acquire(&lock)

```
    }
    printf("Second line (hello=%d)\n", hello);    ←   lock_release (&lock)
    return 0;
}
```

Add in the necessary code to the above problem to make it work correctly.

3.7 SpaceX Problems

Consider this program.

```
pthread_mutex_t lock = PTHREAD_MUTEX_INITIALIZER;
pthread_cond_t cv = PTHREAD_COND_INITIALIZER;

int n = 3;

void* counter(void* arg) {
  pthread_mutex_lock(&lock);
  for (n = 3; n > 0; n--)         ← n = 0
    printf("%d\n", n);
  pthread_cond_signal(&cv);
  pthread_mutex_unlock(&lock);
}

void* announcer(void* arg) {
  while (n != 0) {
    pthread_mutex_lock(&lock);
    pthread_cond_wait(&cv, &lock);    stuck
    pthread_mutex_unlock(&lock);
  }
  printf("FALCON HEAVY TOUCH DOWN!\n");
}

int main() {
  pthread_t t1, t2;
  pthread_create(&t1, NULL, counter, NULL);
  pthread_create(&t2, NULL, announcer, NULL);
  pthread_join(t1, NULL);
  pthread_join(t2, NULL);
  return 0;
}
```

What is wrong with this code?