

Operating Systems, PhD Qualifying Exam, Spring 2015

- This is 2-hour closed-book test. Calculator is allowed.

1. [20 points] Count how many different outputs the following program can print. Assume `fork()` and `wait()` function calls are always successful. `fork()`, on success, returns the child's PID (process ID) in the parent, and returns 0 in the child. `wait()`, on success, returns the terminated child's PID.

```
main() {
    if (fork() == 0) {
        if (fork() == 0) {
            printf("4");
            exit(0);
        } else {
            pid_t pid; int status;
            if ((pid = wait(&status)) > 0) {
                printf("3");
            }
        }
    } else {
        if (fork() == 0) {
            printf("2");
            exit(0);
        }
        printf("1");
    }
    printf("0");
}
```

2. [20 points] Given the following mix of jobs, job lengths, and arrival times, find a preemptive schedule that gives the minimal total turnaround time of all jobs. Justify why there is no better schedule than your answer. The turnaround time of a job is the time at which the job completes minus the time at which the job arrived in the system.

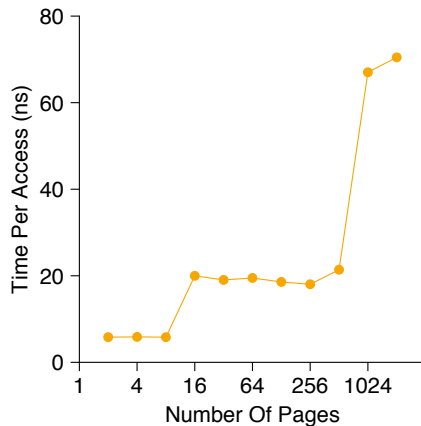
Job	Job Length	Arrival Time
A	80	0
B	30	10
C	35	15
D	40	20

3. [20 points] Explain the difference between race condition, deadlock, livelock, and starvation.

4. [20 points] A system engineer writes the following program in order to measure the size and cost (access latency) of TLB. The program accesses a very large integer array, $A[i]$, and reports the average of array access time at the end. `get_time()` function returns the current time in nano seconds and has no memory accesses.

```
int jump = PAGESIZE / sizeof(int);
for (int i = 0; i < NUMPAGES * jump; i += jump) {
    start_time = get_time();
    A[i] = 0;
    total_access_time += get_time() - start_time;
}
```

- (a) [10 points] Explain why the program is useful to measure the size and cost of TLB.
- (b) [10 points] A system engineer ran the program by successively doubling `NUMPAGES` from 2 to 2048 and obtained the following graph, which shows the measured access time per access in y-axis according to the `NUMPAGES` value in x-axis. Describe the TLB organization in terms of size and cost.



5. [20 points] Virtualization needs an efficient scheme of shared resource management for high resource utilization. Memory is one of resources shared by multiple *guest* operating systems (OS) and managed by the virtual memory system running on *host* OS. When a guest OS needs a new page, the guest OS asks the host OS for a new page and the host OS allocates a page for the guest OS in the physical memory. If different virtual machines run the same code for the OS, many pages for the guest OS can be shared in the physical memory. Unfortunately, the host OS does not know that the page requested by the guest OS is the same as the page that exists in the memory (i.e., previously requested by another guest OS). Thus, the host OS will allocate a new page in the memory, which results in the waste of memory resource. Design an efficient scheme that enables page sharing among guest operating systems.