Elec 428 - Spring 2006
Homework 1

All homework assignments and projects in this course are covered by the Honor Code. You may work on the homework problems together but all written work must be exclusively your own (do not look at anyone else's written work or at solutions from previous years). Late homework assignments lose credit at the rate of 10% per day.

Due date: Wednesday, January 25

1. Suppose that an interactive system is supporting 80 simultaneous users with 10-second think times and a system throughput of 6 interactions/second.

   (a) What is the response time of the system?

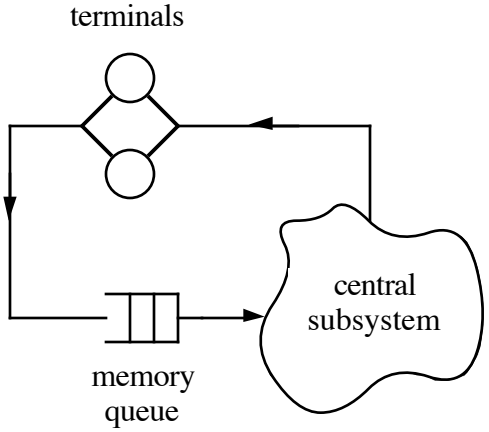
   (b) Suppose the service demands of the workload evolve over time so that system throughput drops to 50% of its former value. Assuming that there are still 80 users with 10-second think times, what would their response time be?

   (c) How do you account for the fact that the response time in (b) is more than twice the response time in (a)?

2. Consider a system modeled as in the accompanying figure. A user request submitted to the system must queue for memory and may begin processing in the central subsystem only when it has obtained a memory partition.

   (a) If there are 100 active users with 20-second think times and system response time (the sum of memory queueing and central subsystem residence times) is 10 seconds, how many customers are competing for memory on the average?

   (b) If the average memory queueing time is 8 seconds, what is the average number of customers loaded in memory?



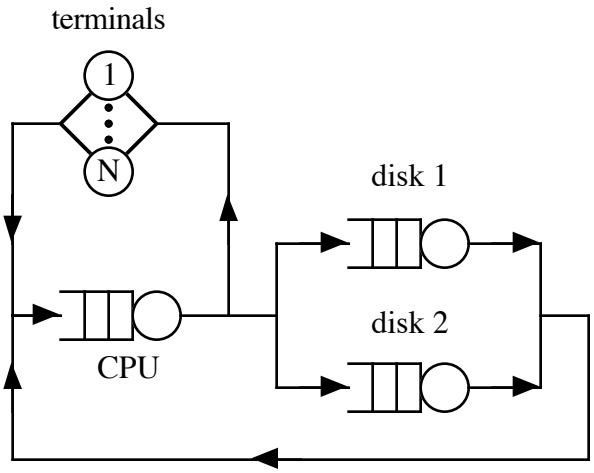
3. (a) Software monitoring of an interactive system shows a CPU utilization of 75%, a 3-second CPU service demand for each interaction (no disks), a response time of 15 seconds, and 10 active users. What is the average think time of these users?

(b) An interactive system with 80 active terminals shows an average think time of 12 seconds. On the average, each interaction causes 15 paging disk accesses. If the service time per paging disk access is 30 ms and this disk is 60% busy, what is the average system response time?

4. In a one-hour observation interval, a particular disk was found to be busy for 35 minutes. If it is known that each job requires 400 accesses to that disk on average and that the average service time per access is 12 milliseconds, what is the system throughput (in jobs/second)?

5. (a) A large database computer system consists of several processors and a disk unit with dozens of disks that supports simultaneous independent accesses to each disk without interference. The system is observed over a period of time when the workload consisted of 20 independent interactive users, each with an average think time 20 seconds. Its disk unit was observed to serve 100 requests per system response time (i.e., each user required 100 disk accesses on average between think times) over the observation period and sustained a combined access rate of 25 requests per second. What is the system's throughput and response time?"

(b) A system consists of 50 terminals ($N = 50$), a CPU, and two disks, as shown in the figure below.



The system was observed for a period of 10,000 seconds and the following data was collected. The average response time that an interactive job experienced was 5 seconds, each job visited the CPU 10 times on the average between think times, and there were 200,000 service completions at the CPU. What was the average total number of jobs at the CPU and disks (i.e., not at a terminal)?

6. An interactive, multiprogrammed computer system consists of 500 users, n CPUs, and m disks. User think times are 10 seconds. When a user finishes "thinking," she sends a job to a CPU, choosing a CPU randomly with equal probability. A job completing its visit to a CPU either returns to the user, with probability 0.1, or goes to a disk. If a job goes to a disk, it chooses a disk randomly with equal probability. When a job completes service at a disk, it goes to a CPU, again choosing a CPU randomly with equal probability. The service demand per visit at a CPU and a disk are $S_{cpu} = 0.01$ seconds and $S_{disk} = 0.02$ seconds, respectively. A disk costs twice as much as a CPU.

(a) Is it possible to achieve a system throughput of at least 100 jobs per second?

(b) What is the minimum possible system throughput with 10 users? 100 users? 1000 users?

(c) What would be the effect on system throughput if a load-balancing strategy were used in selecting CPUs and disks instead of fixed probabilities for given values of n and m?

(d) What is the most cost effective way to configure the system (how many disks and how many CPUs) so that the maximum possible system throughput is at least 40 jobs per second?