Exam 2
CPSC 614
Spring 2001

Name_____

Student ID_____

This exam is 50 minutes, closed book. Show your work for partial credit.

You may find the following formulas for a M/M/m queuing system useful. A M/M/m system has exponentially distributed (Poisson, $C = 1$) arrival and service times, one queue, m servers, server utilization u, arrival rate r, service time T_{ser}, and system time T_{sys}.

$$u = r\frac{T_{ser}}{m},\ T_{sys} = T_{ser}\left(1 + \frac{u}{m(1-u)}\right)$$

You might find these formulas for disk seek time useful. The fallacy equation and its coefficients are:

$$SeekTime(Dist) = a\sqrt{Dist - 1} + b(Dist - 1) + c$$

$$a = \frac{-10Time_{\min} + 15Time_{avg} - 5Time_{\max}}{3\sqrt{NumberCylinders}},\ b = \frac{7Time_{\min} - 15Time_{avg} + 8Time_{\max}}{3NumberCylinders},\ c = Time_{\min}$$

1. (20pts) The probability of N or more tasks in an M/M/1 system is u^N for server utilization u. You are designing a disk controller for a new disk. The disk has an exponentially-distributed service time with an average of 10 ms. The disk I/O requests are exponentially distributed with an average of 50 per second. Assume the I/O request queue will be implemented as a memory in the disk controller. How much memory in terms of request entries is required for the queue, in order for the chance of losing a request due to queue overflow to be less than one request per million?

2. (20pts) A RAID 1 system uses mirrored disks. Each disk holds an identical copy of the file system. The disadvantage of this is that the file system only has the capacity of one disk, no matter how many disks are mirrored together. However the performance advantage of RAID 1 is that all the disks can service read requests in parallel, and any request can go to any disk. Assume your file system is storing read-only data, such as compiler libraries or application binaries, so all requests are read I/Os. Assume also that all requests go to a RAID controller, which then farms out the requests to the first idle disk. Assume the RAID controller overhead is 1 ms. The read I/O request rate is 100 reads per second, and each disk has an average service time of 10 ms (rotation+seek+transfer), both exponentially distributed. What is the fewest number of disks required to achieve an average I/O service time of less than 15 ms?

3. Assume an array of four disks attached to one controller in a RAID 0 configuration. Each disk has 16 surfaces, 6526 tracks/surface, 171 sectors/track, 512-byte sectors, rotates at 10,033 RPM, Time-min is 0.98 ms, Time-avg is 7.71 ms, and Time-max is 18.2 ms. Use 1500 cylinders as the distance in the fallacy equation. The time to switch between surfaces is the same as the time to move the arm one track. In the disk array the spindles are synchronized, so that sector 0 of track 0 of every disk is under the head at the same time. The data is "striped" across all four disks, so four consecutive sectors on a single disk system will be spread one sector per disk in the array. In other words, sectors 0, 1, 2, and 3 of track 0 on a single disk will be stored as sector 0 of track 0 on disks 1, 2, 3 and 4. The delay of the disk controller is 2 ms per I/O request.

(a) (20 pts) Assuming an M/M/1 system, what is the maximum number of I/Os per second and megabytes per second, assuming 2 KB I/O requests.

(b) (20 pts) What is the maximum number of I/Os per second and megabytes per second assuming a single disk, but all sectors in an I/O request are sequential on the disk?

4. (20pts) Given the following data cache/memory system: 95% cache hit rate, block size of 2 words, fetch an entire block on a miss, write an entire block if dirty, the CPU generates memory references at the rate of 10^9 words/s (already taking into account stalls), 25% of the memory references are writes, the memory system can support 10^9 words/s of reads or writes, the memory bus reads or writes a single word at a time (the memory system cannot read or write two words at once), 30% of cache blocks have been modified (are dirty), and the cache uses write allocate on a write miss. What percentage of the memory system bandwidth is used for a write back cache? State any reasonable assumptions.