COMP9334 Capacity Planning
Assignment, Session 1, 2018

March 13, 2018

Instructions

(1) There are 3 questions in this assignment. Answer all questions.

(2) The total marks for this assignment is 15 marks.

(3) In answering the questions, it is important for you to show us your intermediate steps and tell us what arguments you have made to obtain the results. You need to note that both the intermediate steps and the arguments carry marks. Please note that we are **not** just interested in whether you can get the final numerical answer right, but we are **more** interested to find out whether you understand the subject matter. We do that by looking at your intermediate steps and the arguments that you have made to obtain the answer. Thus, if you can show us the perfect intermediate steps and the in-between arguments but get the numerical values wrong for some reason, we will still award you marks for having understood the subject matter.

If you use a computer program to perform any part of your work, you **must** submit the program or you lose marks for the steps.

(4) The submission deadline is 11:00pm Wed 11 April 2018. Late submission will cap the maximum mark that you receive. Submissions after 11:00pm on 16 April will no longer be accepted.

(5) Your submission should consist of:

  (a) A report describing the solution to the problems. This report can be typewritten or a scan of handwritten pages. This report must be in pdf format and must be named assignment.pdf. The submission system will only accept the name assignment.pdf.

  (b) One or more computer programs if you use them to solve the problems numerically. You should use tar/zip/rar to archive all the computer programs into one file with the name supp.tar, supp.zip or supp.rar. The submission system will only accept these names. The report must refer to the programs so that we know which program is used for which part.

(6) Submission can be made via the course website.

Question 1 (2 marks)

An interactive computer system consists of a CPU and a disk. The system was monitored for 60 minutes and the following measurements were taken:

Number of completed jobs	1267
Number of CPU accesses	2,178
Number of disk accesses	2,412
CPU busy time	2,929 seconds
Disk busy time	2,765 seconds

Answer the following questions.

(a) Determine the service demand for each device of the system.

(b) Use bottleneck analysis to determine the asymptotic bound on the system throughput when there are 20 active terminals and the think time per job is 14 seconds.

Note: If you use a computer program to derive your numerical answers, you must include your computer program in your submission. Do not forget to show us your steps to obtain your answer.

Question 2 (7 marks)

A company has a support centre to deal with internal queries from its employees. The support centre currently has 4 staff. The centre has an automatic dispatcher to direct the calls to the support staff. The dispatcher currently has a queue that can hold 2 calls. However, there are no queueing facilities at the staff's terminals. The queueing network at the support centre is depicted in Figure 1.
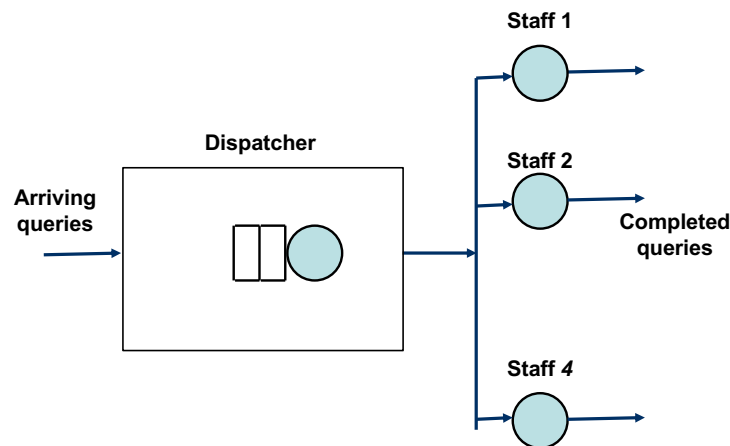

Figure 1: Figure for Question 2.

The voice-over-IP record shows that the centre is getting on average 15 queries per hour. The arrivals can be modelled by using Poisson distribution.

The record also shows that each support staff can complete on average 3 queries per hour. The amount of time required by each query is exponentially distributed.

When a query arrives at the dispatcher, it will accept the query if the dispatcher queue is not full, otherwise the query will be rejected. If a query is accepted and the queue is not empty, the query will be placed at the end of the queue. If a query is accepted and the queue is empty, then the query will be placed in the queue if all staff are busy, otherwise it will be sent to an idling staff. A query will leave the system after its processing is completed. Whenever a staff becomes idle, he/she will take the query from the front of the queue if there is one.

Answer the following questions:

(a) Formulate a continuous-time Markov chain for a system similar to that described above with 4 staff and n waiting slots. Your formulation should include the definition of the states and the transition rates between states. Note that we ask you to use n waiting slots because you will be varying the number of waiting slots in later parts of this question.

(b) Write down the balance equations for the continuous-time Markov chain that you have formulated.

(c) Derive expressions for the steady state probabilities of the continuous-time Markov chain that you have formulated.

(d) For the current configuration, i.e. for $n = 2$, determine:

   (i) The probability that an arriving query will be rejected. Let us denote the result of this by x.

   (ii) The mean waiting time of an accepted query in the queue.

(e) Assuming that you are the manager of the support centre and you think the current rejection rate is too high. You decide to add more waiting slots at the dispatcher to try to reduce the rejection rate. Determine the blocking probability if you add 5, 10, 15 and 20 waiting slots.

(f) Explain why there is little drop in blocking probability after adding 10 waiting slots. What should you do to reduce the blocking probability?

Note: If you use a computer program to derive your numerical answers, you must include your computer program in your submission. Do not forget to show us your steps to obtain your answer.
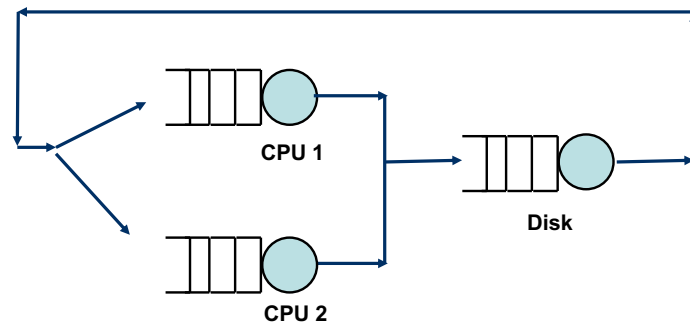
Question 3 (6 marks)

Figure 2: Figure for Question 3.

Consider the computer system shown in Figure 2. The system consists of three devices: a disk and 2 CPUs. Each device is modelled as a server and a queue. The system is at peak load and there are four (4) jobs circulating in the system at all times. During each round that a job circulates the system, the job requires processing from *one* of the CPUs and then followed by the disk. Assuming that:

- The processing time required by each job per visit to the disk is exponentially distributed with mean 200 milli-seconds.

- The two CPUs have different mean processing times. The mean processing times for CPU1 and CPU2 are, respectively, 200 and 400 milli-seconds. Both processing time distributions are assumed to be exponential.

- After a job has left the disk, it will proceed to receive processing at one of the CPUs immediately. In an attempt to utilise the faster CPU (i.e. CPU1), the choice of the CPU depends on the number of jobs at CPU2 at the time when a job leaves the disk. The following job assignment strategy is employed:

  (a) If at the time a job leaves the disk, the total number of jobs at CPU2 is less than 2, then there is an equal probability for that job to go to CPU1 or CPU2.

  (b) If at the time a job leaves the disk, the total number of jobs at CPU2 is 2 or more, then that job will to go to CPU1.

Answer the following questions.

(a) Let the states be the following 3-tuple:

(number of users in the CPU1, number of users in CPU2, number of users in the disk),

formulate a continuous-time Markov chain for this computer system. Your formulation should include (1) a list of states; (2) the transition rates between the states.

(b) Write down the balance equations for the continuous-time Markov chain that you have formulated in Part (a).

(c) What are the steady state probabilities for each state?

(d) What is the throughput of the system?

(e) What is the mean number of jobs in CPU1?

(f) What is the mean response time of CPU1?

Note: If you use a computer program to derive your numerical answers, you must include your computer program in your submission. Do not forget to show us your steps to obtain your answer.

$- - -- $ End of assignment $ - - --$