SOFT 437
Quiz #2
February 26, 2015

Do not turn this page until the quiz officially begins.

STUDENT NUMBER ____ ____ ____ ____ ____ ____ ____ ____

Please do not write your name anywhere on this quiz. I recommend writing your student number at the top of each page.

You are expected to behave considerately towards your fellow students. Noisy or disruptive behavior will not be tolerated. **Turn your phone off.** If you finish early, you may quietly leave the room after handing in your quiz. **However, nobody will be permitted to leave during the last 10 minutes of the quiz.**

If you leave the quiz early, do not stand around outside the exam room talking. This is extremely distracting for students still working.

Academic dishonesty will not be tolerated. Keep your eyes on your own paper or the blackboard at the front of the room. You may not refer to other notes or books during the quiz. You may use a calculator during the quiz.

Please try to write your answers in the space provided. If you need to write your answers elsewhere, please indicate clearly where on the quiz your answers are to be found.

You have 50 **minutes** to complete the quiz. Good luck!

Question 1	/14
Question 2	/6
Question 3	/10
Question 4	/10
Total	/40

Student ID:

Cheat sheet:

$X_i = X_0 \times V_i$ $\quad U_i = X_i \times S_i$ $\quad RT_i = \frac{S_i}{1-U_i}$, $N_i = X_i \times RT_i$ $\quad D_i = V_i \times S_i$

$$D_i = \frac{U_i}{X_0}$$

X_0 =System throughput, X_i, U_i =Throughput and Utilization of device i, V_i = Number of visits, D_i= Service demands

Question 1 **(14 Marks)**

A database server with one CPU and two disks receives requests at a rate of 10,800 requests per hour. Each request needs 200 ms of CPU and performs 5 I/Os on disk 1 and 3 I/Os on disk 2 on average. Each I/O takes 15 ms.

1) What is the average response time per request?
2) What is the average throughput of the server?
3) What are the utilizations of the CPU, disk 1 and disk 2?
4) What is the average number of requests at the server?
5) What is the maximum theoretical arrival rate of requests sustained by this server?

Answer for Q1:

2)
$X0 = \lambda = 10{,}800/3600 = 3$ request/sec

3)
$D_{cpu} = 0.2$ sec
$D_{disk1} = V1 * S_{disk1} = 5 * 0.015 = 0.075$ sec
$D_{disk2} = V2 * S_{disk2} = 3 * 0.015 = 0.045$ sec
$U_{CPU} = D_{cpu} * X0 = 0.2 * 3 = 0.6 = 60\%$
$U_{disk1} = D_{disk1} * X0 = 0.075 * 3 = 22.5\%$
$U_{disk2} = D_{disk2} * X0 = 0.045 * 3 = 13.5\%$

1) (note that you need the previous calculations)
Residence times can be computed as:
$RT_{cpu} = D_{cpu}/(1-U_{cpu}) = 0.2/(1-0.6) = 0.5$ sec
$RT_{disk1} = D_{disk1}/(1- U_{disk1}) = 0.075/(1-0.225) = 0.097$ sec
$RT_{disk2} = D_{disk2}/(1 - U_{disk2}) = 0.045/(1-0.135) = 0.052$ sec

Total response time $RT = 0.5 + 0.097 + 0.052 = 0.649$ sec

4)
$N = RT * X0 = 0.649 * 3 = 1.947$ requests

5)
The maximum arrival rate is $1/\max\{0.2, 0.075, 0.045\} = 5$ requests/sec

Student ID:

Question 2 **(6 Marks)**

A transaction server is planned to be part of the infrastructure of a Web service. It is estimated that the service should support 800 concurrent users. The average think time is 20 seconds, which means that after receiving a reply from the Web service, each user waits on average 20 seconds to submit a new request. Every service request accesses the transaction server 15 times. The average response time of the request is 4 seconds. What is the throughput for the transaction server?

Answer for Q2:

Concurrent users: M = 800
Think time: Z = 20 sec
RT = 4

M = X0 * (Z + RT) → X0 = M/(Z+RT) = 800/(20+4) = 33.33 request/sec

The throughput for the transaction server: X_{ts} = 33.33 * 15 = 500 requests/sec

Student ID:

Question 3 **(10 Marks)**

1) Explain the major components that determine the adequate capacity of an IT infrastructure.
2) Describe the business evolution planning, customer behavior planning, and IT resource planning.
3) Explain how these three aspects are used together in capacity planning.

Answer for Q3:

1- Adequate capacity for an e-business site is a function of the three following elements
   - Service level agreement (SLA)
   - Technology
   - Cost Constraints

   [refer to the lecture for more details]

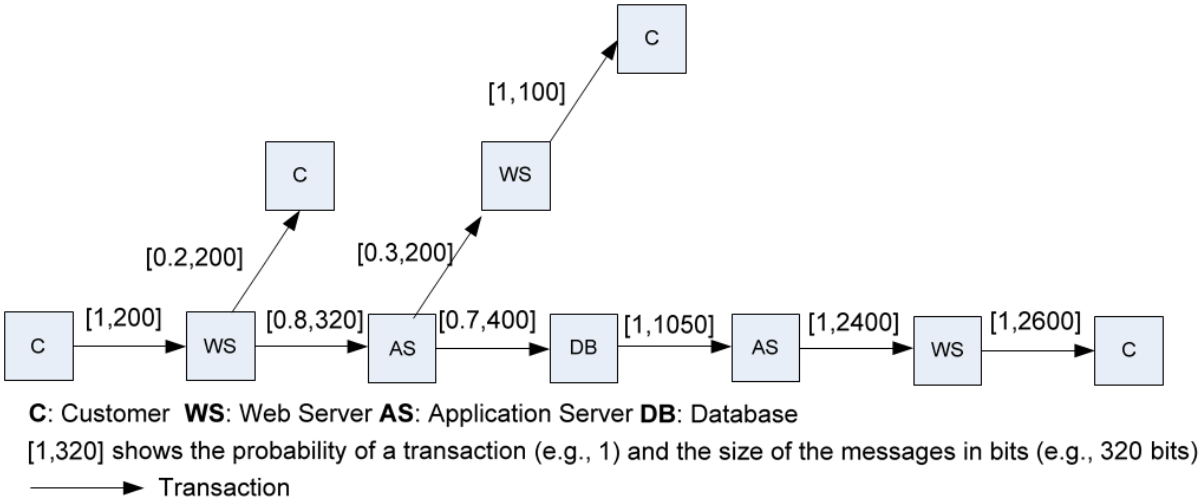
2- Business evolution planning is the studying of how business objectives could affect the site's performance and required capacity. Good planning makes preparations to accommodate the potential expansion and changes in the system behavior to maintain performance objectives.
   Customer behavior planning is the characterization of customer behavior as a result of implementing the business evolution.
   IT resource planning is the generation of a description of the IT infrastructure as well as a workload description. Description of the IT infrastructure includes hardware, software, networking, payment services, etc.

3- Performance engineers must study the business evolution plans and evaluate how that impacts the user behavior. Then, they capture (or model) this change through simulations or investigating system logs to determine what are the required resources to maintain/meet the performance objective.

Student ID:

Question 4 **(10 Marks)**

The workload characterization is the process of precisely describing, in a qualitative and quantitative manner, the global workload of an e-business site. A Client/Server Interaction Diagrams (CSID) is built for an end-to-end business function, as illustrated the diagram below. Use the CSID to answer the following questions.



1) What is the probability that the database server is used during the execution of this e-business function? **(2 Marks)**
2) What is the number of average times the application server is used? **(2 Marks)**
3) The average number of bytes that cross the LAN for the execution of this e-business function? **(2 Marks)**
4) Suppose that the protocol byte overhead is 15% and that the LAN is a 100Mbps Ethernet. Assume that the effective bandwidth of the network, after contention due to network traffic is taken into account, is 80 Mbps. What is the maximum number of execution of this e-business function per unit time? **(2 Marks)**
5) Suppose that the application server and the database server are in the same machine. Recalculate 4) **(2 Marks)**

Student ID:

Answer for Q4:

1) P(DB) = 1 * 0.8 * 0.7 = 0.56

2) The average time of uses for application server is:
   1* (1 *0.8) + 2 * (1*0.8*0.7) = 1.92

2) 1 * 0.2 * (200 + 200) + 1 * 0.8 *0.3*1* (200 + 320 + 200 + 100) + 1 *0.8*0.7(200+ 320+ 400 + 1050 + 2400 + 2600) = 4180

4) N = $(80 * 10^6)/(4180 * (1+ 15\%))$ = 16642

5) Average message size=
   1 * 0.2 * (200 + 200) + 1 * 0.8 *0.3*1* (200 + 100) + 1*0.8*0.7(200+ 2600) = 1720

   N = $(80 * 10^6)/(1720 * (1+ 15\%))$ = 40445