

Midterm Exam

رقم الشعبة: 1

الرقم التسلسلي:

الاسم:

Instructions: Time **60** minutes. Open book and notes exam. No electronics. Please answer all problems in the space provided and limit your answer to the space provided. **No questions are allowed.**

<Good Luck>

Q1. Assume that you want to evaluate a memory system composed of a DRAM memory module and a bus that interconnects this memory with the processor. Also assume that the specifications of both the memory module and the bus are known to you.

[10 marks]

a) Among the three main evaluation techniques, which one do you use? State your assumptions and justification for the selection you make.

Assuming that I don't have the system to perform measurements, I would use analytical modeling because it is sufficient for this simple system and is not expensive.

b) How would you validate the results of your selected technique?

Using measurements on a similar system.

If no such a system is available, I must use simulation.

c) What performance metrics would you use in your study?

Peak performance, sustained bandwidth, response time, error probability.

d) Recall that there are two definitions for the response time. What definition would you use and why?

I would use Definition 1 if the processor supports critical word first and early restart.

Otherwise, I would use Definition 2.

e) Assuming that you want to use a workload in this study, what type of workload would you use and why?

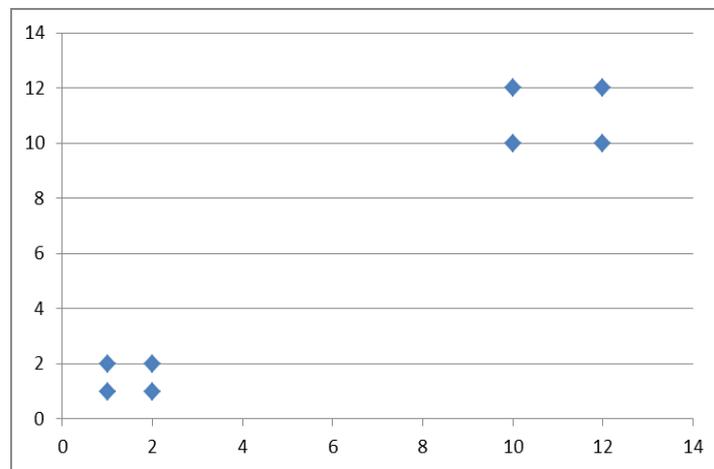
Microbenchmark: synthetic workload that focuses on exercising the memory system.

Q2. The workload of a computer system was monitored and two metrics were measured (m_1 and m_2) for eight workload instances shown in the table below. Using the technique you think is the best in this case, characterize this workload.

[5 marks]

Instance	m_1	m_2
1	1	1
2	2	1
3	1	2
4	2	2
5	10	12
6	12	10
7	10	10
8	12	12

When the data is plotted on a scatter graph (as show below), we notice that there are two distinct clusters.



So we use clustering and report the centers of the two clusters:

Cluster 1: (1.5, 1.5)

Cluster 2: (11, 11)

Q3. Assume that you have the following code, instrument this code so that you can thoroughly monitor its performance, i.e., rewrite it with the added instrumentation code.

[5 marks]

```
main() {
    int size;
    printf("Enter size: ");
    scanf("%d", &size);
    int x[] = init_array(size);
    calc_array(x);
    print_results(x);
}
```

```
main() {
    int size;
    printf("Enter size: ");
    scanf("%d", &size);
    printf("Calling init_array with size %d at time %d/n",
        size, time());
    int x[] = init_array(size);
    printf("Calling calc_array at time %d/n", time());
    calc_array(x);
    printf("Calling print_results at time %d/n", time());
    print_results(x);
    printf("Return from print_results at time %d/n", time());
}
```

Q4. Given the following performance of two systems on three benchmarks,

Benchmark	System A (tps)	System B (tps)
I	1000	2000
II	2000	1600
III	60	30

Using ratio games, how the makers of System A would advertise the advantage of their system?

[5 marks]

As the metric is higher is better, we use System B as reference.

Benchmark	System A	System B
I	0.5	1
II	1.25	1
III	2.0	1
Average	1.25	1

Therefore, makers of System A would advertise that their system has a relative performance of 25% higher than System B using these three benchmarks.

Q5. Using SimPack, write a simulator to find the utilization of a system composed of one server. Assume that the inter-arrival time has an exponential distribution of mean = 60 seconds and the service time has an exponential distribution of mean = 50 seconds. The simulator should simulate the arrivals of 1000 jobs.

[5 marks]

```
#include <stdio.h>
#include "queuing.h"

#define NEW_ARRIVAL      1
#define REQUEST_SERVER  2
#define RELEASE_SERVER   3

TOKEN a_token;

int server,          /* server facility descriptor */
    njs = 1000;     /* no. of jobs to simulate */

main() {
    int event
    init_simpack(LINKED);
    server = create_facility("Server", 1);
    schedule(NEW_ARRIVAL, 0, a_token);

    while (njs) {
        next_event(&event, &a_token);

        switch(event) {
        case NEW_ARRIVAL:
            schedule(REQUEST_SERVER, 0, a_token);
            update_arrivals();
            schedule(NEW_ARRIVAL, (TIME) expntl(60), a_token);
            break;

        case REQUEST_SERVER:
            if (request(server, a_token, j) == FREE)
                schedule(RELEASE_SERVER, (TIME) expntl(50), a_token);
            break;

        case RELEASE_SERVER:
            update_completions();
            njs--;
            break;

        }
    }
    report_stats();
}
```