VU VRIJE UNIVERSITEIT AMSTERDAM

This is a written exam for the course "*Performance of Networked Systems*"
Lecturers: prof.dr. R.D. van der Mei and dr. T. Kielmann
Date: Friday, March 28, 2014

Rules for the exam:

1. Allowed material: This is a closed-book exam. You are not allowed to use all kinds of written material or your laptop, and electronic communication during the exam is strongly prohibited.

2. Calculation of end grade for the course: the end grade for the course is built up in two parts: homework assignments and a written exam.

   - *Homework assignments*: The average of the two grades counts to 40% of the final grade.
   - *Written exam*: for this written exam you get a grade between 1 and 10. This grade will count for the remaining 60% of the final grade.
   - *Final grade*: the final grade is calculated as the weighted average of the grade for the written exam on the one hand, and the average homework grade on the other hand, with the restriction that the grade for the written exam must be at least 4.0.

3. Credits: This written exam consists of four questions (A, B, C and D), each of which consists of a number of sub-questions. The maximum number of points you can get is distributed as follows amongst the sub-questions:

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**total**
A	4	4	4	4	4				**20**
B	4	3	3	3	4	4	**4**		**25**
C	5	4	4	4	5	3			**25**
D	6	6	4	4					**20**

Good luck

QUESTION A: Some basics

A.1 What is a Poisson process, and why do Poisson processes provide a natural way to model randomly occurring events?

A.2 A discrete random variable *N* is said to have a Poisson distribution with mean λ if for k=0,1,2,…

$$\Pr\{N = k\} = \frac{e^{-\lambda}\lambda^k}{k!}.$$

What is the relation between a Poisson *process* and a Poisson *distribution*?

A.3 The Poisson process is known to have two important properties: (1) the *superposition* property, and (2) the *thinning-out* property. What do these properties mean?

A.4 During the course the M/G/1 processor sharing (PS) model was discussed. For this model it was shown that the expected sojourn time is given by the following expression

$$E[S] = \frac{\beta}{1-\rho},$$

where β is the mean service time and ρ is the load of the system. This result implies that E[S] is insensitive with respect to the service-time distribution. What does that mean?

A.5 Mean Value Analysis (MVA) is a powerful way to calculate performance metrics for closed queueing networks. Explain in words what the basic idea behind MVA is (no mathematical notation is required).

QUESTION B: The simplified database server case

We consider the database server example that was used in the homework assignment and the textbook, but with a single disk (see Figure 1 below). To this end, we consider a computer system with a single CPU and a single disk to support a database server. Users remotely access the server and typically login, perform some database transactions, and logout.

It is assumed that exactly K customers are logged onto in the system at any time. That is, as soon as one user completes the requested transaction and exits the system, another customer is waiting to log on. The additional assumptions are:

- Each transaction alternates between the CPU and using the disk.
- The amount of CPU time required by a transaction is exponentially distributed with a mean of 10 seconds.
- The time to access a file on the disk is exponentially distributed with mean 20 seconds.

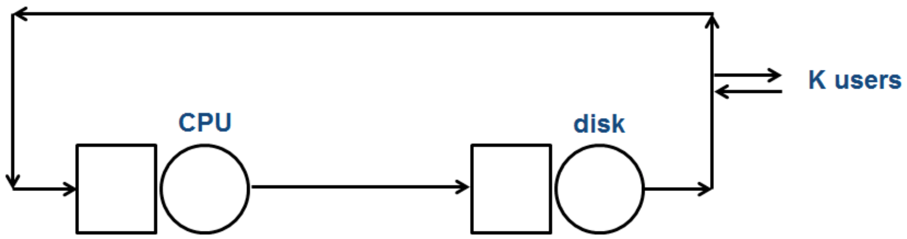

Figure 1. The database server example with single disk.

B.1 Consider the simple case K = 1. That is, there is only a single customer in the system. For this case, define a continuous-time Markov chain: (1) *How are the state defined?*, (2) *What are the possible states?*, and (3) *What are the state-transition rates.*

B.2 Define the state probabilities and formulate the balance equations.

B.3 Calculate the state probabilities by solving the balance equations.

Next, use these results to answer the following questions

B.4 What response time can the typical user expect? What is the utilization of both of the system resources? What is the throughput of the system?

Hint: do not only give numbers, use proper argumentation!

Let us now take the case K = 2. That is, there are two users in the system all the time.

B.5 Define a continuous-time Markov chain: (1) *How are the state defined?*, (2) *What are the possible states?*, and (3) *What are the state-transition rates.*

B.6 Formulate and solve the balance equations.

B.7 What response time can the typical user expect? What is the utilization of both of the system resources? What is the throughput of the system?

QUESTION C: Operational Laws

C.1 Assume a corporate intranet Web site (like vunet.vu.nl). Sketch a queueing network (with the web site as a black box) where in total M users, at their computers, are alternatingly thinking or waiting for a response from vunet.vu.nl. vunet's response time should be denoted as R, the thinking time as Z, and vunet's throughput as X0.

C.2 We distinguish between three types of resources in queueing networks. In C.1, to which type does the web site belong? Which type are the users? What is the third type of resource that is not in the queueing network from C.1? Give an example of how this resource type could be used to extend the network from C.1.

C.3 Give an intuitive explanation of Little's law.

C.4 Apply Little's law to compute the average number of users waiting for vunet (N') and the average number of users thinking (M').

C.5 By combining the formulas for M' and N', derive a formula that computes R as a function of M, X0, and Z. What is the name of this formula? (e.g., which operational law is this?)

C.6 Assume that vunet has 1000 users at the same time, with Z=10sec, and R<=1sec. How many requests per second does vunet have to be able to process?

QUESTION D: Measurements and Models

D.1 For assessing the performance of a system, one can either measure the system itself, or simulate its behaviour, or use an analytical model. For each of these three alternatives, explain the advantages and disadvantages.

D.2 For measurements of a system, you need a workload under which the system is observed. Explain the difference between natural and artificial workload models. Give examples for both (two each).

D.3 For a sample of measurements, you can compute the *coefficient of variation.* Give an intuitive explanation of this coefficient and its use for analysing the measurement results.

D.4 For a sample of measurements, give an intuitive explanation of *confidence intervals* as a tool to describe the measured property.