Norwegian University of Science and Technology
Department of Telematics

Contact during exam [Faglig kontakt under eksamen]:
Bjarne E. Helvik (92667)

EXAM IN COURSE [EKSAMEN I EMNE]
TTM4120 Dependable Systems [Pålitelige systemer]

Tuesday [Tirsdag] 2010-05-25
09:00 – 13:00

The English version starts on page 2.

Den norske bokmålsutgaven starter på side 10.

Hjelpemidler:
D - No printed or handwritten material is allowed. Predefined simple calculator [Ingen trykte eller håndskrevne hjelpemidler tillatt. Forhåndsbestemt enkel kalkulator]

Sensur 2010-06-16

English version[1]

This exam deals with some dependability issues related to a server park. It has a hardware architecture as illustrated in Figure 1. The servers are divided into two groups, each having its individual independent power supply. The internal communication system and the access to the Internet are duplicated in two separate sides, so a bus and/or a router on one side may fail without disturbing the operation of the system. The servers, the routers, the internal communication buses and the power supplies fail with constant failure rates λ_s, λ_r, λ_b and λ_p respectively. Failures occur independently of each other. The other elements indicated may be assumed to be fault free.

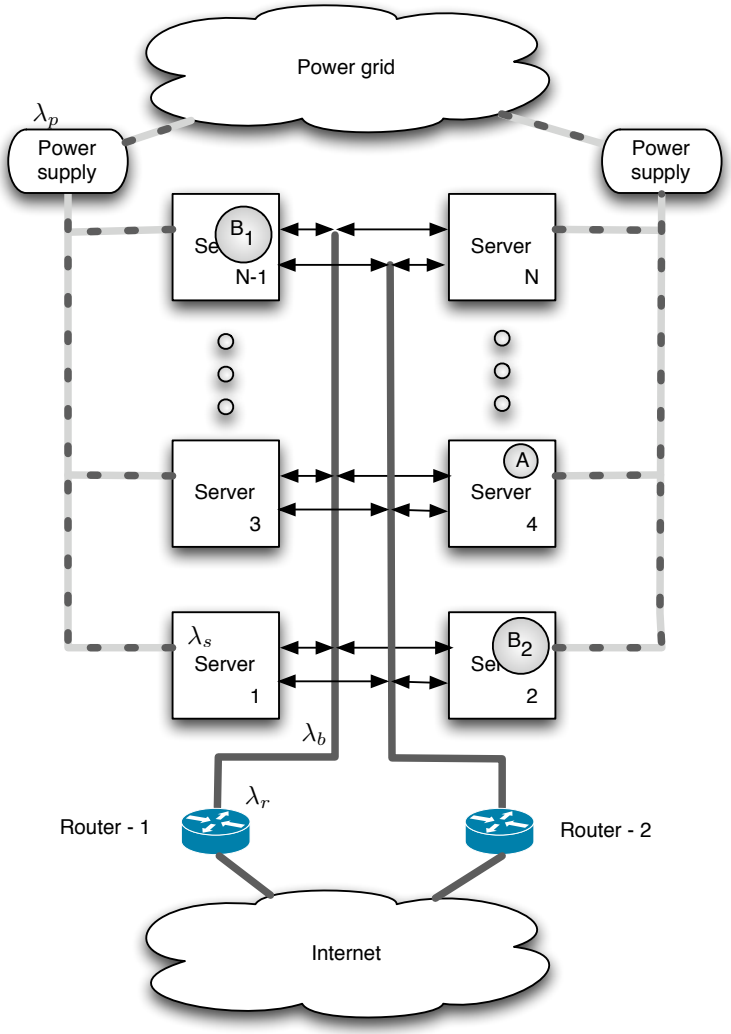

Figure 1: Hardware architecture of the server park.

[1] In case of divergence between the English and the Norwegian version, the English version prevails.

a) Assume that the system starts and all system entities are working at $t = 0$. What is the failure intensity, $z_A(0)$ at this point in time of a process A running on server 4? Motivate the answer. Give the formal definitions of failure intensity, $z(t)$, and failure rate, $\lambda(t)$. What are the relation between the failure intensity $z_A(0)$ and the failure rate $\lambda_A(0)$ at the beginning of system operation?

Informally the failure intensity, $z_A(0)$ is the expectation (probability) per time unit of a failure occurring immediately after start. In the outlined system, this may happen if power supply 2 or the server hosting process A fails. Other failures are either irrelevant (i.e. other server failures) or will not result in immediate system failure (i.e., the communication system that is duplicated). Hence, $z_A(0) = \lambda_s + \lambda_p$.

Failure rate is the probability per time unit that the system/unit will fail after a time t of uninterrupted operation, i.e. $\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr(t<T\leq t+\Delta t | T>t)}{\Delta t}$ where $T$is the time of uninterrupted operation.

Failure intensity is the expected number of failures per time unit occurring at time t (which is equal to the probability per time unit that the system/unit will fail at time t if there is at most one failure at a time (i.e. regular process)), i.e., $z(t) = \lim_{\Delta t \to 0} \frac{E(N(t+\Delta t)-N(t))}{\Delta t}$ where $N(t)$is the number of failures that has occurred at time t.

It follows from the definition that for our case $z_A(0) = \lambda_A(0)$. (This is in fact the line of reasoning leading to the answer $z_A(0) = \lambda_s + \lambda_p$ above.)

b) If there are no repairs in the system, determine the reliability function $R_A(t)$ of process A. Use this result to show how we may find its failure rate $\lambda_A(t)$. What is the limit of the failure rate when the process has been operational for a long period, i.e., the limit $\lambda_A(\infty)$? It is sufficient that the result is obtained by a direct (physical) argument.

To determine $R_A(t)$ it is easiest to make a block diagram including the units that supports process A, as shown in Figure 2. Note that we study the time to first failure, so reallocation of process A to another server after the failure is not an issue.

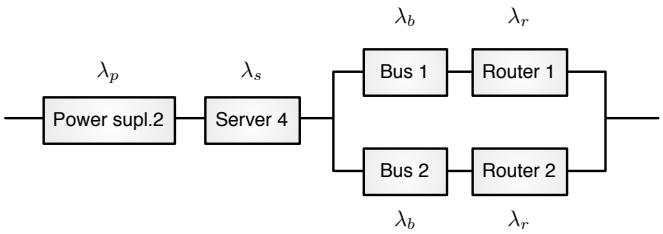

Figure 2: Block diagram showing the support of process A

From this scheme the reliability function is obtained

$$\begin{aligned} R_A(t) &= \exp(-(\lambda_p+\lambda_s)t)(1-(1-\exp(-(\lambda_b+\lambda_r)t))^2) \\ &= \exp(-(\lambda_p+\lambda_s+2(\lambda_b+\lambda_r))t)(2\exp((\lambda_b+\lambda_r)t)-1) \end{aligned}$$

We have that $\lambda_A(t) = \frac{-\frac{d}{dt}R_A(t)}{R_A(t)}$. This may be remember or derived from the formal definition in item a) Expanded this yields ***(nb!** expansion is not required)*

$$\lambda_A(t) = \frac{2\left(-1+e^{t(\lambda_b+\lambda_r)}\right)\lambda_b + \left(-1+2e^{t(\lambda_b+\lambda_r)}\right)\lambda_p + 2e^{t(\lambda_b+\lambda_r)}\lambda_r - 2\lambda_r + 2e^{t(\lambda_b+\lambda_r)}\lambda_s - \lambda_s}{-1+2e^{t(\lambda_b+\lambda_r)}}$$

The limit $\lambda_A(\infty)$ may be obtained by direct argument, i.e., the limiting failure rate is reached when all spare capacity has failed and hence limit $\lambda_A(\infty) = \lambda_p + \lambda_s + \lambda_b + \lambda_r$. It may also be obtained from the result above. (*Not required*: Figure 3 shows the case where $\lambda_p = \lambda_s = \lambda_b = \lambda_r = 1$.

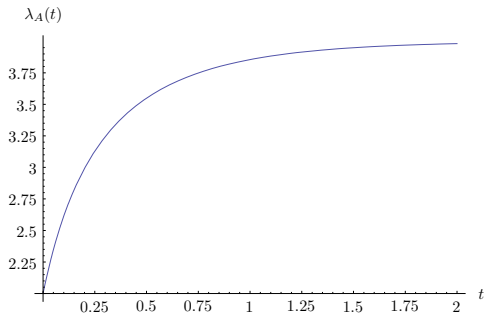

Figure 3: The process A failure rate when all unit failure rates are one (1).

c) Assume that all system entities are repaired independently of each other and have a constant repair time of d_x, where $x, = s, r, b, p$ according to the type of entity. What is the asymptotic availability, A_x, of an entity of type x? We would like to obtain the asymptotic availability of the the server park, A_{park}, when we require that at least K, where $N \geq K > N/2$, out of the the N servers must provide service towards the Internet for the system to be operational. What modelling method may be used? Motivate the answer. Make a dependability model of the server park and obtain an expression for A_{park}. Establish the equations necessary to obtain the mean time between failures of the server park, $\text{MTBF}_{\text{park}}$. Note that it is *not* necessary to solve or reduce the equations.

It is see seen that all failures and repairs are independent. Hence, we may use reliability block schemes and combinatorial methods for the evaluation. (The constant repair times does not pose an obstacle in this context.)

The asymptotic availability of an entity x may generally be obtained from the relation $A_x = \frac{\text{MUT}_x}{\text{MDT}_x+\text{MUT}_x} = (1 + d_x\lambda_x)^{-1}$.

Since $N \geq K > N/2$ both power supplies must work. From Figure 1 it is seen that it is sufficient that one communication branch works, since all servers may interact with the Internet via both. The resulting block scheme is shown in Figure 4

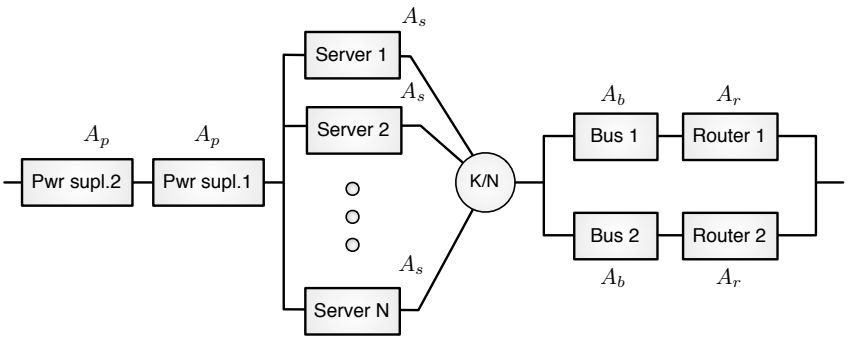

Figure 4: Reliability block scheme of the server park when $N \geq K > N/2$

From this figure, it is directly obtained that $A_{\text{park}} = A_p^2(1-(1-A_bA_r)^2)\sum_{i=K}^{N}\binom{N}{i}A_s^i(1-A_s)^{N-i}$.

For the MTBF, the overall series structure of the subsystems may by used first

$$\text{MTBF}_{\text{park}} = 1/(A_{\text{park}}(\lambda_{\text{power}} + \lambda_{\text{servers}} + \lambda_{\text{comm}}))$$

where the lambdas are the failure intensities of the subsystems when working, i.e. $\lambda_{\text{xxx}} = 1/\text{MUT}_{\text{xxx}}$, $\lambda_{\text{power}} = 2\lambda_p$ and $\lambda_{\text{servers}} = \binom{N}{K}A_s^K(1-A_s)^{N-K} \cdot K\lambda_s/A_{\text{servers}}$, where $A_{\text{servers}} = \sum_{i=K}^{N}\binom{N}{i}A_s^i(1-A_s)^{N-i}$. For the communication system, we first regard the two halves where we have to find the MDT of a half, i.e. $\text{MDT}_{\text{com-half}} = (1/(\lambda_r+\lambda_b)\frac{1-A_rA_b}{A_rA_b}$, which yields $\lambda_{\text{comm}} = \frac{U_{\text{comm}}}{1-U_{\text{comm}}}\frac{2}{\text{MDT}_{\text{com-half}}}$ where $U_{\text{comm}} = (1-A_rA_b)^2$.

On the servers shown in Figure 1 runs a middleware, JaSoS used in the laboratory of the course, supporting the processes. The middleware is not perfect, so a server failure will with probability $1-c$ crash the middleware platform, i.e., processing on all servers stops. The middleware is restarted

again (on the non-failed servers) in a negative exponentially distributed time with expectation d_c. Assume that the other repair times in the system are negative exponentially distributed as well, with expectations d_x, where $x, = s, r, b, p$ according to the type of entity. The active repair times are independent. However, at most one failure is repaired at a time.

Two processes, B_1 and B_2, provides the same service. There is a huge number of clients using this service. Each client interaction involves considerable computations. If these are interrupted because of failures, the intermediate state of client requests may be lost without causing a failure in the service provisioning. However, the final state caused by a request should be maintained and be available for further processing also after a server failure. The processes operate load shared and one may take the load of the other if it fails. They run on server $N-1$ and 2, and will not be reallocated to or restarted on another server if these fails, but they will be restarted on the same server after it is repaired.

d) Taking into account the middleware, what is the failure intensity of process A, regarded in Question a), when all system entities are working?

In the following, the process pair B and the service they provide are regarded. To simplify, assume that:

- all servers but $N-1$ and 2 have zero repair time, i.e., $d_s^{(z)} \rightarrow 0$ for $z = 1, 3, \ldots, N-2, N$ and $d_s^{(z)} > 0$ for $z = 2, N-1$;
- The routers and internal communication do not fail, i.e., $\lambda_r = \lambda_b \rightarrow 0$.
- It is not necessary to expand the state diagram further (include more failure events) when the service of B has failed.
- Unpowered servers will not fail.

Under the above assumptions and the further assumption that the power supplies do not fail, i.e., $\lambda_p \rightarrow 0$, establish a state diagram (Markov model) that may be used to determine the availability of the service provided by the process pair B, $A_B(t)$.

Next, *remove* the assumption that the power supplies does not fail, i.e., $\lambda_p > 0$, and extend the model to include this cause of failure.

In the modelling, identify the operational mode of each state clearly. Indicate whether the service is available or not in each state. Show which part of the model (i.e. which states and transitions) that describes the case without power supply failures.

The failure intensity of process A is $z_A^*(0) = (N-1)(1-c)\lambda_s + \lambda_s + \lambda_p$. See arguments below, is included to serve as a hint.

The coverage issue is dealt with in the lecture notes. The difference is that here, it is sufficient to include the states of the two servers executing the process B replicas, c.f. the immediate repair assumption for the others. The "crux" of the first part of the modelling is to include that the other servers may cause platform failures and thereby failures of the the two regarded, cf. the hint given by the first sub-question. The platform restart will leave the two processor regarded up or down

dependent of their relative fraction of the total number. See Figure 5. In the diagram, it is used that $\mu_x = d_x^{-1}$.

The extension is about dealing with a more complex system with more failure causes, where also the repair limitation (and sequence) becomes relevant. When there is multiple faults in the system, it must be decided which to repair first. In most cases the repairs that will bring the system from failed to working, is given priority. Next, faults that when repaired, will make the system more robust vs. new failures are given priority, i.e. power supply faults before server faults. Furthermore, repairs that have short repair time are given priority over those with longer.

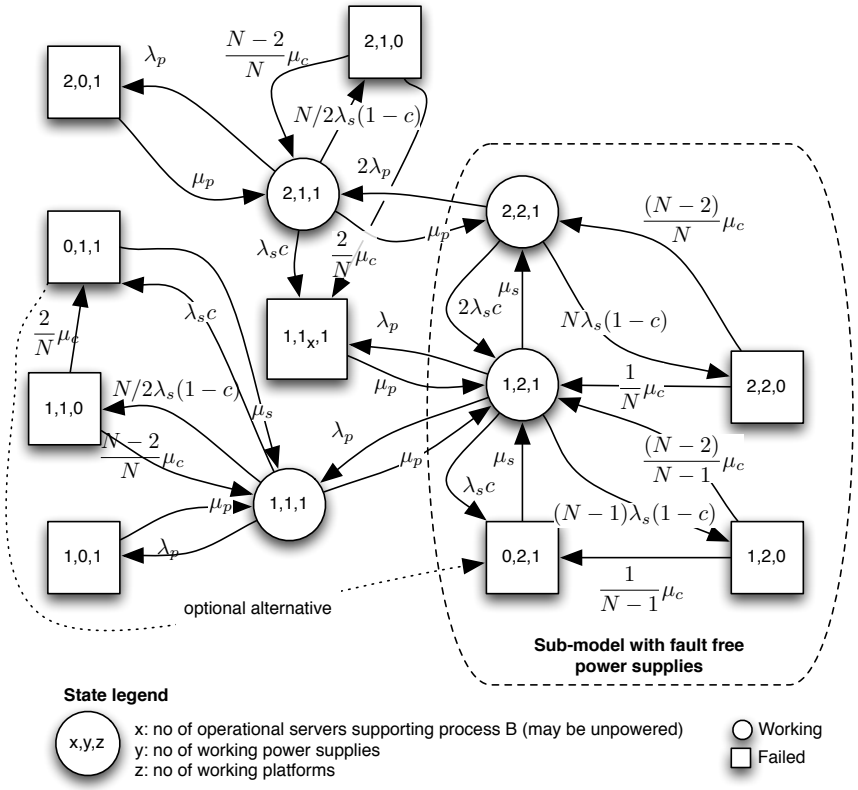

Figure 5: Markov model of the service provided by the process pair B

e) JaSoS (as well as some other group communication systems) distinguish between internal and external group invocations. Explain what is meant by internal and external group invocations, and the rationale for this division. What is the difference between anycast and multicast invocations? Outline briefly how you would apply these operations to make the processes B_1 and B_2 to provide a service tolerant to single server failures.

A group constitute one entity in the system seen from outside the group.

internal group invocations: are invocations on the group by its members

external group invocations: invocations on the group from other parts of the system/its clients.

A consistent view of the group and its (member's) state should be maintained through operations. One way of doing this is to include the clients into the group before interactions with the servers are performed. In this case the same kind of invocations may be used. The clients leave the group after the interaction is finished. This is, however, not efficient/doable when there are a huge number of clients. Hence, the rationale of the division is to enable the group members to have a set of internal operations maintaining a strong consistency within the group, and to have external operations on the group that maintains consistency within the group but allowing a weaker synchronization between the server group and client. The internals of the group are hidden from the clients.

anycast invocations: invocations on one arbitrary member of a group

multicast invocations: simultaneous invocations on all members of a group.

Clients makes an anycast invocation to start the computations, which are performed by one server. When the result is reached, the server makes a multicast (IGMI) to establish a common internal state and returns the result to the client.

The middleware platform, JaSoS, is continuously debugged and improved to increase c, i.e. to improve its reliability. It is hypothesized that this reliability growth may be modelled by the relation $z_p(t) = \alpha + \beta \exp(\delta t)$, where in the above context $z_p(t) = N\lambda_s(1 - c(t))$.

f) Give a short physical interpretation of the three parameters α, β and δ, and indicate their range when we require that the model should be physical and that we in fact have a growth. According to this model, what is the expected number of platform crashes (failures) in the interval $[0, \tau]$, and what is the probability that there is no crashes (failures) in the interval, as a function of the model parameters? We have observed crashes/failures at the times $t_1, t_2, \ldots, t_n$. Establish a set of equations that may be used to estimate the three parameters. What method do you use?

Given: The likelihood function of an inhomogeneous Poisson process with intensity $z(t)$ where events takes place at $x_1, x_2, \cdots, x_m$ is $L(x_1, x_2, \cdots, x_m) = \exp(-\int_0^{x_m} z(t)dt) \cdot \prod_{i=1}^{m} z(x_i)$.

$\alpha \geq 0$ is the lower bound on the the failure intensity, which is reached when $t \to \infty$.

$\beta \geq 0$ is the the increase in the failure intensity above the lower bound observed at the beginning of the improvement period.

$\delta \leq 0$ is determining the time constant $(-\delta)^{-1}$ in the growth, i.e. $1 - e^{-1}$ of the potential growth is completed after $(-\delta)^{-1}$.

The expected number is $E[N(\tau)] = Z(\tau) = \int_0^\tau z_p(t)dt = \alpha\tau + \frac{\beta}{\delta}(\exp(\delta\tau) - 1)$ and since we from the assumptions have that the overall process is an inhomogeneous Poisson process, $\Pr(N(\tau) = 0) = \exp(Z(\tau))$.

The maximum likelihood estimator may be used, i.e. finding the set $\hat{\alpha}$,$\hat{\beta}$, $\hat{\delta}$, which maximizes the likelihood (or log likelihood since log is a monotone function). Hence, we should find the set $\hat{\alpha}$,$\hat{\beta}$, $\hat{\delta}$ that satisfies the set of equations $\frac{\partial}{\partial\alpha}\log(L(t_1, t_2, \ldots, t_n)) = 0$, $\frac{\partial}{\partial\beta}\log(L(t_1, t_2, \ldots, t_n)) = 0$ and $\frac{\partial}{\partial\delta}\log(L(t_1, t_2, \ldots, t_n)) = 0$, and is a global maximum.