Norges teknisk-naturvitenskapelige universitet
Institutt for telematikk

EKSAMEN I EMNE
TTM4120 PÅLITELIGE SYSTEMER

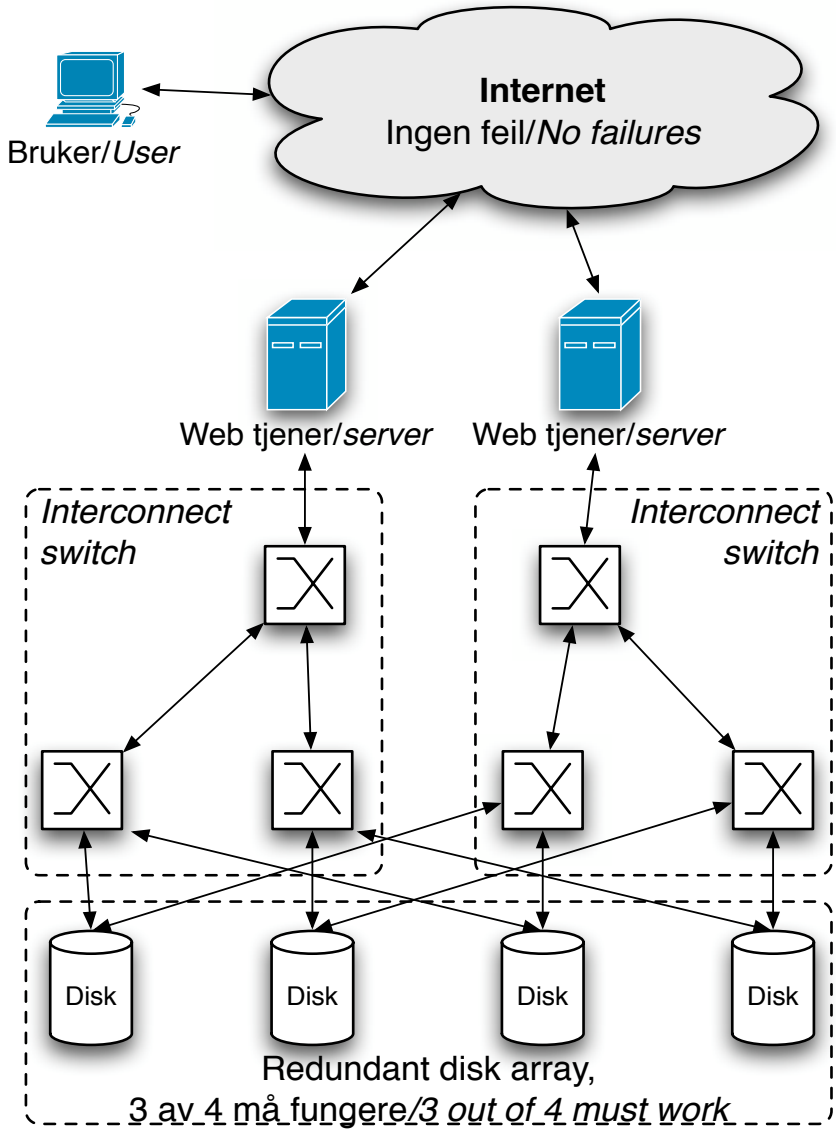
Faglig kontakt under eksamen: Bjarne E. Helvik
Tlf.: 92667

Eksamensdato:	30. mai 2006
Eksamenstid:	4 timer
Studiepoeng:	7,5 Vt
Tillatte hjelpemidler:	D
Språkform:	
Antall sider bokmål:	3
Antall sider nynorsk:	1
Antall sider engelsk:	3
Antall sider vedlegg:	0
Sensurdato[1]:	uke 25 - 2006

1. Merk! Studentene må primært gjøre seg kjent med sensur ved å oppsøke sensuroppslagene. Evt. telefoner om sensur må rettes til sensurtelefonene. Eksamenskontoret vil ikke kunne svare på slike telefoner.

BOKMÅLSUTGAVE

Vi har et lastdelt web tjenersystem som vist i Figur 1 nedenfor. Det består av to uavhengige tjenermaskiner. En bruker logger på en av disse og er pålogget samme maskin til tjenesteleveransen er gjennomført. Hvilken tjenermaskin som blir brukt avgjøres av en ideell lastdelingsfunksjon i nettet. Denne funksjonen vil kun dirigere nye tjenesteforespørsler til arbeidende maskiner, og en maskin er tilstrekkelig til å håndtere hele trafikklasten. For leveranse av tjenesten benytter og oppdaterer de to maskinene samme data. Disse er lagret i et array av redundante disker. Data kan leses og skrives så lenge minst tre av de fire diskene fungerer. Diskene aksesseres via en "interconnect switch" med struktur som vist i figuren. Brukerne, Internet og lenkene mellom de ulike systemelementene kan antas å være feilfrie.



Figur 1 Skisse av et lastdelt tjenersystem

Det stilles krav til *tilgjengeligheten* til tjenesten og til *funksjonssikkerheten* til tjenesten.

a) Definer (forklar hva som menes med) disse to pålitelighetegenskapene[1] og definer et formelt mål for hver av dem.

Anta at alle systemelementene feiler og repareres uavhengig av hverandre. Følgende størrelser antas kjent

	Utilgjengeligheten	Feilrate
Tjenermaskin	U_t	λ_t
Svitsjeelement (1x2) i "interconnect switch"	U_s	λ_s
Disk	U_d	λ_d

b) Finn et uttrykk for utilgjengeligheten til tjenesten som leveres fra systemet.

c) Kan pålitelighetsblokkskjema benyttes for å finne den funksjonssannsynlighet som en bruker pålogget en tjenestemaskin opplever? Begrunn svaret. Anta at diskene (disk-delsystemet) ikke feiler. Finn da et uttrykk for funksjonssannsynlighet som en bruker av tjenesten opplever.

d) Funksjonssikkerheten for tjenesten er for dårlig. Skisser en endring av systemløsningen som kan benyttes for å forbedre denne. En ønsker ikke å skifte ut tjener maskinvaren eller å anskaffe ytterligere maskinvare.

I det etterfølgende antas svitsjeelementene og diskene å være feilfrie. Dvs. vi ser bort fra disse.

De to tjenermaskinene oppgraderes gjentatte ganger for å gi økt funksjonalitet. Oppgraderingsintensiteten er θ, dvs. at dersom tjenerparet ikke allerede er under oppgradering, så vil en slik startes med en sannsynlighet $\theta \cdot \Delta t + o(\Delta t)$ i løpet av et kort tidsintervall Δt. En tjenermaskin som *ikke* er under oppgradering har tilgjengelighet og feilrate som gitt ovenfor. Anta at feilavhjelpingstidene (reparasjonstidene) er negativt eksponensialfordelt. Kall parameteren i denne μ_t. Oppgraderingene foregår i fire faser etterfulgt av en ordinær driftsfase som følger:

Fase 1: Idet en oppgraderingskommando kommer, går maskinene inn i en *forberedingsfase*. Varigheten av denne fasen er negativt eksponensialfordelt med parameter β. Det er ordinær trafikkhåndtering samt feil kan inntreffe og reparasjon foretas i denne fasen. (Forberedelsene foregår kun når begge maskiner er oppe.)

Fase 2: I denne fasen *tas en maskin ut av drift og oppgraderes*. Varigheten av oppgraderingen dersom den forløper uten feil er negativt eksponensialfordelt med parameter γ. Maskinen under oppgradering har en feilrate som øker til α. Dersom den feiler vil den ha samme fordeling av feilavhjelpingstid (reparasjonstid) som ved andre feil, dvs. negativt eksponensialfordelt med

1. Engelsk: Dependability attributes

parameter μ_t. Feiler den andre maskinen, dvs. den som fører last, vil begge maskinene kjøre en feilavhjelping. Ved feil i denne fasen startes fra fase 1 igjen.

Fase 3: Begge maskinene har *ordinær trafikkhåndtering med hhv. ny og gammel versjon*, og er i en *forberedingsfase* som i fase 1. Ellers som i fase 1.

Fase 4: Den andre *maskinen tas ut av drift og oppgraderes*. Ellers som i fase 2 med det unntaket at ved feil i denne fasen feilavhjelpes systemet tilbake til starten av fase 3.

Fase 5: *Ordinær drift* av systemet.

e) Tegn et tilstandsdiagram som kan benyttes til pålitelighetsanalyse av tjenerparet med oppgradering som beskrevet ovenfor. Angi tydelig konfigurasjonen i de ulike tilstandene, hvilken fase du assosierer tilstandene med og hvorvidt de er arbeidende- eller feiltilstander.

f) Finn et uttrykk for sannsynligheten for at en oppgradering gjennomføres uten at det inntreffer feil i løpet av prosessen (gjennomløpet av de fire fasene). Hva er forventet tid en slik feilfri oppgradering tar?

Transisjonsmatrisen for tjenersystemet kan avledes av tilstandsdiagrammet i punkt e). Den har følgende struktur:

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{21} & 0 & 0 & \lambda_{51} \\ \lambda_{12} & \lambda_{22} & 0 & 0 & 0 \\ 0 & \lambda_{23} & \lambda_{33} & \lambda_{43} & 0 \\ 0 & 0 & \lambda_{34} & \lambda_{44} & 0 \\ 0 & 0 & 0 & \lambda_{45} & \lambda_{55} \end{bmatrix} \quad (1)$$

hvor λ_{ij} er en submatrise som angir transisjonene mellom tilstandene i fase i og tilstandene i fase j. "0" angir submatriser hvor alle elementer er 0.

g) Etabler submatrisen som svarer til $\begin{bmatrix} \lambda_{11} & \lambda_{21} \\ \lambda_{12} & \lambda_{22} \end{bmatrix}$.

h) Etabler et fullstendig ligningssett for å beregne asymptotisk utilgjengeligheten til webtjenesten når den oppgraderes som beskrevet ovenfor. Transisjonsmatrisen Λ kan antas kjent. (Merk, ligningssettet skal ikke løses.)

i) Kall tiden fra en oppgadering starter til den er gjennomført for T_o. Vis hvordan du med utgangspunkt i (1) kan finne sannsynligheten for at denne tiden er lengre enn t, dvs. $P(T_o > t)$, når du antar at begge maskinene er arbeidende idet en oppgradering starter. Etabler et fullstendig ligningssett.

NYNORSKUTGÅVE

For ikkje å få ulike semantikk i bokmål og nynorsk oppgåvetekst så er inga eige nynorsk oppgåvetekst laga. Sjå i staden bokmåloppgåve på side 2-5. Dette etter avtale med “nynorsk studentane”.

Ved problem ta kontakt med faglærar.

ENGLISH VERSION

We have a load shared web server system as illustrated in Figur 1 on Page 2. It has two independent servers. A user loggs on one of these and stays logged on this server until the delivery of the service is completed. Which server that will be used in each case, is decided by an ideal load sharing function in the network. This function will distribute a new service request only to working servers, and one server has sufficient capacity to handle the entire load. For the delivery of the service, the two servers use and update a common set of data. These data are stored in an array of redundant disks. The data may be read and written as long as at least three out of the four disks are working. The disks are accessed via an interconnect switch with a structure as illustrated in the figure. The users, Internet and the links between the different system element are assumed to be fault free.

There are *unavailability* and *reliability* requirements of the provided service.

a) Define (explain what is meant by) these two dependability attributes and define a formal measure for each of them.

Assume that all system elements are failing and are repaired independently of each other. The following quantities are assumed to be known

	Unavailability	Failure rate
Server	U_t	λ_t
Switching element (1x2) in interconnect switch	U_s	λ_s
Disk	U_d	λ_d

b) Find an expression of the unavailability of the service delivered by the system.

c) May a reliability block diagram be used to find the reliability function a user logged on to a server experiences? Justify the answer. Assume that the disks (the disk subsystem) do not fail, and derive an expression for the reliability function a user of the system experiences.

d) The reliability function of the service is poor and does not meet the requirements. Sketch a change in the system design, which may be used to improve the reliability. The server hardware shall not be replaced and additional hardware shall not be procured.

In the subsequent, it is assumed the switching elements and the disks are fault free, i.e. these are not regarded.

The two servers are upgraded on several occasions to yield an increased functionality. The intensity of the upgrades are θ, i.e., if the pair of servers are not already under upgrade, a new upgrade will start with probability $\theta \cdot \Delta t + o(\Delta t)$ during a short interval Δt. A server that is not under upgrade, has an unavailability and failure rate as stated above. Assume that the restoration times after a failure (repair times) are negatively exponentially distributed with parameter μ_t. The upgrades take place in four phases followed by an ordinary operational phase as follows:

Phase 1: When a upgrade is started, the servers enters an *initialization phase*. The duration of this phase is negatively exponentially distributed with parameter β. During this phase, there are ordinary handling of the traffic load, and failures may occur and restoration (repair) take place. (The initialization takes place only when both servers are working.)

Phase 2: In this phase, *one server is taken out of operation and upgraded*. The duration of the upgrade, if it takes place without failure, is negatively exponentially distributed with parameter γ. The server under upgrade has a failure rate which increases to α. If it fails, it will have the same distribution of the restoration time as with other failures, i.e. negatively exponentially distributed with parameter μ_t. If the other server fails, i.e., that one that carries the traffic load, both servers are restored. If a failure occurs in this phase, the upgrade restarts from Phase 1.

Phase 3: Both servers *carry traffic load with a new and old version* of the software and are *under initialization* as in Phase 1. Otherwise as in Phase 1.

Phase 4: The *other server is taken out of operation and upgraded*. Otherwise as in Phase 2, with the exception that when failures occur in this phase, the system are restored back to the start of Phase 3.

Phase 5: *Ordinary operation* of the system.

e) Make a Markov model (state diagram) that may be used for dependability analysis of the pair of servers with the upgrade procedure outlined above. Indicate clearly the system configuration in the various states and whether they are working (up) or failed (down) states. Do also indicate which of the five phases each of the states belong to.

f) Derive an expression for the probability that an upgrade are completed without any failures during the procedure (during the four phases). What is the expected time such a “fault free” upgrade takes?

The transition matrix of the server system may be derived from the state diagram in item e). It has the following structure:

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{21} & 0 & 0 & \lambda_{51} \\ \lambda_{12} & \lambda_{22} & 0 & 0 & 0 \\ 0 & \lambda_{23} & \lambda_{33} & \lambda_{43} & 0 \\ 0 & 0 & \lambda_{34} & \lambda_{44} & 0 \\ 0 & 0 & 0 & \lambda_{45} & \lambda_{55} \end{bmatrix} \tag{2}$$

where λ_{ij} is a sub-matrix that gives the transitions between the states of Phase i and the states of Phase j. “0” indicates sub-matrixes where all elements are 0.

g) Establish the sub-matrix corresponding to $\begin{bmatrix} \lambda_{11} & \lambda_{21} \\ \lambda_{12} & \lambda_{22} \end{bmatrix}$.

h) Establish a complete set of equations, which may be used to determine the asymptotic unavailability of the web-service described above. The transition matrix Λ may be assumed to be known. (Note, the set of equations shall not be solved.)

i) Denote the time from an upgrade is started and until it is completed by T_o. Show how you from (2) may obtain the probability that this time is longer than t, i.e., $P(T_o > t)$, when you assume that both servers are working when the upgrade starts. Establish a complete set of equations.