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]

Monday [Mandag] 2011-05-30
09:00 – 13:00

The English version starts on page 2.

Den norske bokmålsutgaven starter på side 6.

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

Sensur 2011-06-20

English version[1]

A company delivering services over the Internet has a fault tolerant ICT system at its operational site. A sketch of the system is shown in Figure 1. It consists of two routers that are homed to different ASes in the Internet, and both of them are connected to two computing clusters. The asymptotic availability of routers, links and computing clusters are A_r, A_l and A_c. From a performance point of view, it is sufficient that one router and one cluster is working for the site to deliver adequate service.

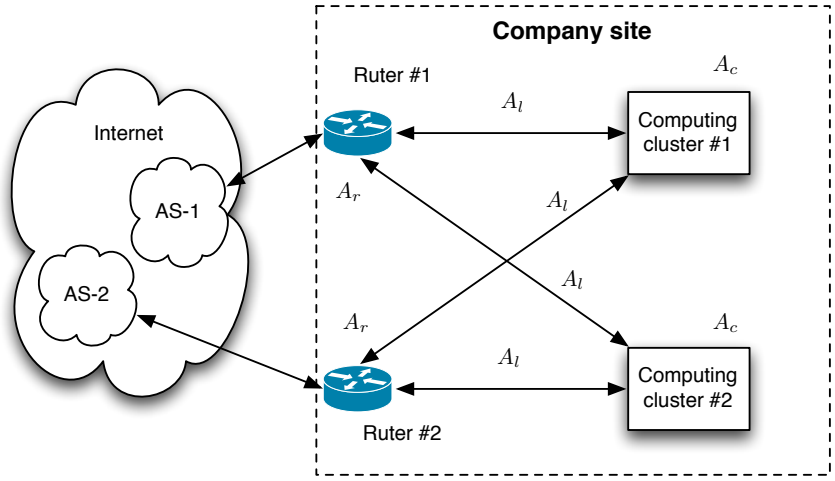

Figure 1: Operational site with fault tolerant ICT system.[Operasjonssenter med feiltolerant IKT system.]

The system has been operational for some time, and a dependability analysis is launched to improve its reliability and availability.

The routers have failed several times. Router #1 has failed eight times. The time to first failure and between the subsequent failures are 0.046, 0.027, 8.051, 0.265, 0.075, 0.024, 12.865 and 1.670 days. Down times after failures are negligible and the router is as new after it is restored/repaired after a failure, i.e the time to first failure and the times between failures have the same statistical properties.

a) Make a plot of the empirical reliability function based on these observations. What is the observed MTFF=MTBF in this case? What is the probability that the router will have a time between failures that is longer than the MTBF? Make the estimate directly from the observations.

b) If a failure process Poissonian (e.g that of Router #1), what is the probability of observing a time between failures that is longer than the MTBF? Motivate the answer. If you should fit the failure data by a common distribution, which one would you choose? Motivate the answer,

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

i	0	1	2	3	4	5	6	7	8	9
Φ_i	{ }	{r1}	{r2}	{l11}	{l12}	{l21}	{l22}	{c1}	{c2}	{r1, r2}
i	10	11	12	13	14	15	16	17	18	19
Φ_i	{r1, l11}	{r1, l12}	{r1, l21}	{r1, l22}	{r1, c1}	{r1, c2}	{r2, l11}	{r2, l12}	{r2, l21}	{r2, l22}
i	20	21	22	23	24	25	26	27	28	29
Φ_i	{r2, c1}	{r2, c2}	{l11, l12}	{l11, l21}	{l11, l22}	{l11, c1}	{l11, c2}	{l12, l21}	{l12, l22}	{l12, c1}
i	30	31	32	33	34	35	36			
Φ_i	{l12, c2}	{l21, l22}	{l21, c1}	{l21, c2}	{l22, c1}	{l22, c2}	{c1, c2}			

Table 1: Indexing of all operational modes with two or fewer failed network elements. ri indicates that router i has failed, ci indicates that cluster i has failed and lij indicates that the link between router i and cluster j has failed. (The third row is intentionally left blank for your use.) [Indeksering av alle operative modi med to eller færre feilte nettelementer. ri indikerer at router i har feilet, ci indikerer at klynge i har feilet og lij indikerer at lenken mellom ruter i og klynge j har feilet. (Den tredje raden er med hensikt latt være tom og er for ev. bruk under eksamen.)]

give the reliability function for the distribution and indicate ranges of parameter(s) if relevant. Let say you select a distribution with pdf f and parameters a and b, i.e., $f(t; a, b)$, *outline* the method you would use to obtain an estimator of the parameters. (It is required that you identify the method, present its principle and give a formal description of how the principle is applied. It is not expected that you derive a specific estimator or estimate parameters from the data.)

c) What is the criteria for a system have a series parallel structure? Does the site have a series parallel structure? Assume that routers, links and computer clusters fail and are repaired independently of each other. What technique can be used to enable a reliability block diagram analysis of the site? Use this technique and find the availability of the service from the site.

The ICT system will have a number of operational modes depending on which network elements are working or not. A subset of these is listed in Table 1.

d) Show how an upper and lower bound of the availability of the services from the site may be obtained by considering only operational modes with two or fewer failed network elements. Obtain the expression for the upper bound with the same assumptions as in c) above.

In the remaining questions we concentrate on a computing cluster. It consists of n processors operating in load sharing. $k \leq n$ out of these are needed to provide the service with sufficient quality. In the configuration in Figure 1, one cluster is active taking the entire load, the other one is a stand-by. The intensity of permanent, i.e. hardware, failures are negligible and we may assume that *all* failures are

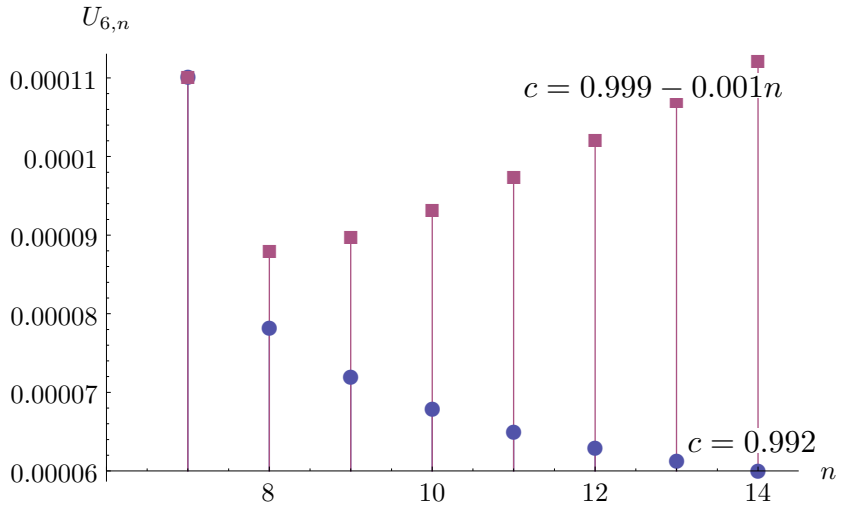

Figure 2: Unavailability of a computing cluster for a range of spare capacities and two coverage models. [Utilgjengelighet av en datamaskinklynge for ulik reservekapasitet og to dekningsmodeller.]

transient or due to logical faults. When a processor fails, it is restarted in a negative exponential time with expectation μ^{-1}. Processors restart independently of each other and all restarts are i.i.d. The failure intensity per processor is λ. When a processor fails, with probability $(1 - c)$, *all* processors in the same cluster stop and have to be restarted.

e) Establish a Markov model (state diagram) which may be used to determine the asymptotic availability, A_c, of the computing cluster. Let p_i be the probability that the cluster has i failed processors. Establish a set of equations that may be used to find $p_i, i = 0, \ldots, n$ and A_c for general n. (The equations shall not be solved.)

In Figure 2 the unavailabilities of the computer cluster, i.e., $U_{6,n} = 1 - A_c$, is obtained from the equations found in e) for $n = 7, \ldots, 14$, in the case where $k = 6$ when $\mu = 1$ and $\lambda = 0.001$. The two following models of the coverage factor c are considered:

1. $c = 0.992$, i.e., a constant probability that a processor failure does not cause the entire cluster to fail.

2. $c = 0.999 - 0.001n$, i.e., the probability that a processor failure does not cause the entire cluster to fail decreases linearly with the number of processors in the cluster.

Six processors are sufficient to provide the services from the site. In the configuration in Figure 1 each computing cluster has $n = 7$ processors, i.e. one processor provides the spare capacity in the cluster in case of failure. A new configuration is considered, where the two clusters are combined and all processors are operated as one load-shared cluster of $n = 14$ processors. The objective of this is to be able to tolerate more failures.

f) What is the number of failed processors that can be tolerated in the two configurations? Will the configuration of Figure 1 or the new single cluster configuration provide the highest availability? Answer the question for both the case with coverage model 1 and the case with coverage model 2 presented above.

g) Assume that we have ideal fault handling, i.e., $c = 1$. Are there any dependencies between the processors in this case? Under this assumption, will the configuration of Figure 1 or the new single cluster configuration provide the higher availability? Hint: it is not necessary to perform numerical calculations to answer the question.

Norsk bokmål utgave[2]

Et selskap som leverer tjenester over Internett har et feiltolerant IKT-system ved sitt operasjonssenter. En skisse av systemet er vist i figur 1. Det består av to rutere som er tilknyttet forskjellige ASer i Internett (eng: multihomed), og begge er koblet til to databehandlingsklynger. Den asymptotiske tilgjengeligheten av rutere, lenker og databehandling klynger erA_r, A_l og A_c. Fra et ytelsessynspunkt, er det tilstrekkelig at en ruter og en klynge virker for at det skal leveres tilfredsstillende service.

Systemet har vært operativt i noen tid, og en pålitelighetsanalyse er startet for å forbedre påliteligheten og tilgjengeligheten.

Rutere har feilet flere ganger. Router #1 har feilet åtte ganger. Tiden til første feil og tidene mellom etterfølgende feil er 0.046, 0.027, 8.051, 0.265, 0.075, 0.024, 12.865 og 1.670 dager. Nedetiden etter feil er ubetydelig, og ruteren er som ny etter at den er restartet/reparert, dvs. tiden til første feil, og intervallene mellom feil har de samme statistiske egenskaper.

a) Lag et plott av den empiriske funksjonsannsynligheten (eng: reliability function) basert på disse observasjonene. Hva er den observerte MTFF = MTBF i dette tilfellet? Hva er sannsynligheten for at ruteren vil ha en tid mellom feil som er lengre enn MTBF? Baser estimatet direkte på feilobservasjonene.

b) Hvis en feilprosess, f.eks. den fra router #1, er en Poisson prosess, hva er sannsynligheten for å observere en tid mellom feil som er lengre enn MTBF for prosessen? Grunngi svaret. Hvis du skal tilpasse observasjonene av tid til/mellom feil til en vanlig brukt sannsynlighetsfordeling, hvilken ville du benytte? Grunngi svaret, angi funksjonsannsynligheten (eng: reliability function) for fordelingen og indikerer verdiområdet til parametere(ne) hvis det er relevant. La oss si du velger en distribusjon med sannsynlighetstetthet (pdf) f og parametre a and b, dvs. $f(t; a, b)$, skisser/beskriv den metoden du vil bruke til å finne en estimator for parametrene. (Det kreves at du identifiserer metoden, presenterer de/det prinsipp den baserer seg på og gir en formell beskrivelse av hvordan prinsippet(ene) anvendes. Det forventes ikke at du utleder en estimator for en spesifikk funksjon eller estimerer parametere fra data.)

c) Hva er kriteriene for et system har en serie-parallell struktur? Har system på operasjonssenteret serie-parallell struktur? Anta at rutere, lenker og datamaskinklynger feiler og repareres/idriftsettes uavhengig av hverandre. Hvilken teknikk kan anvendes for å gjennomføre en pålitelighetblokkdiagramanalyse av operasjonsenteret? Bruk denne teknikken til å finne tilgjengeligheten til tjenestene fra senteret.

IKT-systemet vil ha en rekke operasjonelle modi avhengig av hvilke nettelementer som virker eller er feilet. Et utvalg av disse er vist i tabell 1.

[2]I tilfelle uoverensstemmelse mellom den engelske og norske utgaven, er det den engelske som er gjeldende. Engelske betegnelser anvendes hvor ingen norsk oversettelse ble funnet.

d) Vis hvordan en øvre og en nedre grense av tilgjengeligheten av tjenestene fra senteret kan finnes ved å ta hensyn til kun operasjonelle modi med to eller færre feilte nettelementer. Finn et uttrykk for øvre grense med de samme forutsetninger som i c) ovenfor.

I de etterfølgende spørsmålene konsentrerer vi oss om en datamaskinklynge. Den består av n prosessorer som opererer i lastdeling. $k \leq n$ av disse er nødvendig for å tilby tjenester med tilstrekkelig kvalitet. I konfigurasjonen i figur 1, er en klynge aktivt og tar hele lasten, mens den andre er reserve (eng: stand-by). Intensiteten av permanente feil, dvs. maskinvarefeil, er ubetydelig, og vi kan anta at alle feil er transiente eller skyldes logiske feil. Når en prosessor feiler, blir den restartet i løpet av en negativt eksponensial fordelt tid med forventning μ^{-1}. Prosessorer restarter uavhengig av hverandre og alle restarter er i.i.d. Feilintensiteten per prosessor er λ. Når en prosessor feiler, vil med sannsynlighet $(1 - c)$ *alle* prosessorer i samme klynge stoppe og må restartes.

e) Etabler en Markovmodell (tilstandsdiagram) som kan benyttes til å bestemme asymptotisk tilgjengelighet, A_c, av klyngen. La p_i være sannsynligheten for at klyngen har i feilte prosessorer. Etabler et sett av likninger som kan brukes til å finne $p_i, i = 0, \ldots, n$ og A_c for en generell n. (Likningene skal ikke løses.)

I figur 2 er utilgjengelighetene til klyngen, dvs. $U_{6,n} = 1 - A_c$, bestemt fra ligningene funnet i e) for $n = 7, \ldots, 14$, når $k = 6$, $\mu = 1$ og $\lambda = 0.001$. Følgende to modeller av dekningsfaktoren (eng: coverage factor) c er tatt i betraktning:

1. $c = 0.992$, dvs. en konstant sannsynlighet for at en prosessorfeil ikke fører til at hele klyngen stopper.

2. $c = 0.999 - 0.001n$, dvs. sannsynligheten for at en prosessorfeil ikke forårsaker at hele klyngen stopper avtar lineært med antall prosessorer i klyngen.

Seks prosessorer er tilstrekkelig til å levere tjenestene fra operasjonsenteret. I konfigurasjonen i figur 1 har hver klynge $n = 7$ prosessorer, dvs. en prosessor sørger for reservekapasitet i klyngen i tilfelle feil. En ny konfigurasjon er vurdert, hvor de to klyngene slås sammen og alle prosessorer opereres som en lastdelt klynge med $n = 14$ prosessorer. Målsetting med dette er å kunne tolerere flere samtidige feil.

f) Hva er antall feilte prosessorer som kan tolereres i de to konfigurasjonene? Vil konfigurasjonen i figur 1 eller den nye konfigurasjonen med én klynge gi den høyeste tilgjengeligheten? Svar på spørsmålet både for tilfellet med dekningsmodell 1 og tilfellet med dekningsmodell 2 presentert ovenfor.

g) Anta at vi har ideell feil håndtering, dvs. $c = 1$. Er det noen avhengigheter mellom prosessorene i dette tilfellet? Under denne forutsetningen, vil konfigurasjonen i figur 1 eller den med én klynge gi den høyeste tilgjengelighet i dette tilfellet? Hint: det er ikke nødvendig å utføre numeriske beregninger for å svare på spørsmålet.