



EKSAMEN I EMNE
SIE5015 PÅLITELIGHET OG YTELSE MED SIMULERING

Faglig kontakt under eksamen: Poul E. Heegaard
Tlf.: 99286858

Eksamensdato: 13. desember 2002
Eksamenstid: 4 timer
Vekttall: 2,5 Vt
Tillatte hjelpemidler: C¹

Språkform:
Antall sider bokmål: 8
Antall sider nynorsk: 0
Antall sider engelsk: 8 (starter på [side 10](#))
Antall sider vedlegg: 13 (Tillatt formelsamling)

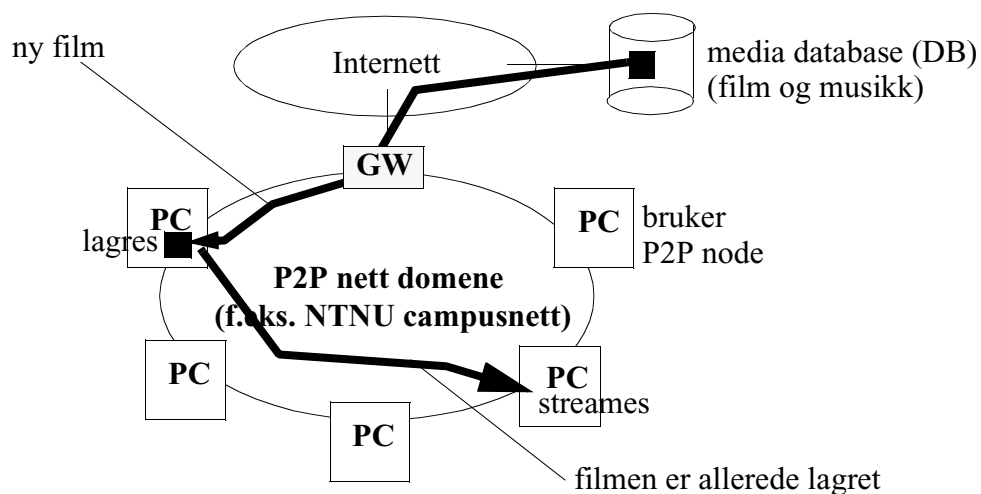
Sensurdato²: uke 2 - 2003

1. Graham Birtwisle: DEMOS - A system for Discrete Event Modelling on Simula. (Personlige annotering i DEMOS bok er tillatt) Formelsamling i fag SIE5015 Pålitelighet og ytelse med simulering. NB! Formelsamlingen er vedlagt på [side 18](#).

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

Internett ble laget for å gjøre det enkelt å dele dokumenter og informasjon. Overføringsmekanismene var enkle og programmene i endessystemene likeså. Etter hvert har antall brukere og mengde informasjon økt eksplodivt. Derfor har det blitt utviklet stadig bedre metoder og programmer for å finne og dele informasjon. Begrepet peer-to-peer kommunikasjon brukes ofte som et samlebegrep på fildelingsmetoder og programmer, f.eks. Napster, Gnutella, Kazaa, directconnect, osv.

I denne oppgaven skal vi ta for oss en fiktiv peer-to-peer tjeneste som er laget for å dele film og musikk. I det etterfølgende bruker vi forkortelsen P2P for peer-to-peer. Tjenesten er laget slik at brukerne får tilgang til å "streame" film og musikk fra andre brukere. Hvis en bruker ikke finner den filmen eller musikken som han leter etter så vil søket sendes til et sett av spesialiserte mediedatabaser som er tilgjengelig via en gateway mot resten av Internett. Fila må i disse tilfellene lagres. Søket går alltid først gjennom all film og musikk lagret hos brukerne i P2P nettet slik at samme film eller plate ikke blir lastet ned over gatewayen flere ganger enn nødvendig. Alle brukere registrerer seg i et register gjeldende for det domenet hvor tjenesten er tilgjengelig. Det er en entydig mapping mellom bruker, dennes PC og en P2P node i P2P nettet. Figuren under viser et slik tenkt domene.

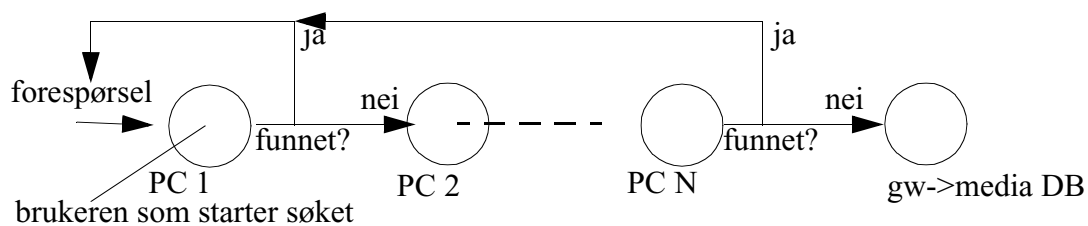


Figur 1 Peer-to-peer (P2P) nett: eksempel med 5 terminaler.

Tjenesten krever registrering og kun et spesifisert utvalg av databaser kan aksesseres. Dette skyldes at det er knyttet betaling og garantier til denne tjenesten. Dette skiller denne fiktive tjenesten fra de fleste andre P2P tjenester i dag. I det etterfølgende vil vi se nærmere på hvilken strategi for nedlasting brukerne bør velge innenfor denne betalingsmodellen, og hvordan garantert kvalitet bør formuleres og beregnes.

Oppgave 1 Ytelsen til søkealgoritmen

Anta nå at en bruker ønsker å se på en film. Først søker han gjennom alle filmer som er lagret hos alle brukere. Søket foregår ved at det sendes en forespørsel til alle brukerne etter tur. Brukeren starter søket hos seg selv. Forespørselen returneres til avsender hvis filmen er funnet, eller videresendes til neste bruker på lista hvis ikke. Hvis forespørselen returneres til avsender uten at filmen er funnet så sendes forespørselen til mediedatabasene som er tilgjengelig for denne tjenesten. Vi antar at filmen eller musikken alltid blir funnet i en av disse databasene.



Figur 2 Søk etter film eller musikk

Vi gjør følgende antakelser

- Under søket skjer det ingen endringer i antall brukere eller filmer lagret av brukerne
 - En spesifikk film er lagret hos en gitt bruker med sannsynlighet p (denne sannsynligheten er lik for alle filmer og alle brukere, også den brukeren som starter søket)
 - τ_s - tiden per søkesteg [antas konstant]
- a) Anta at antall brukere N er uendelig. Ved bruk av søkealgoritmen som beskrevet over: Hva er fordelingen til antall søkesteg, X , i P2P nettet? (merk at hvis brukeren finner filmen hos seg selv så blir det $X = 0$ søkesteg) Hva kalles denne fordelingen? Beregn forventet søketid, $E(T)$, og varians, $Var(T)$, for den beskrevne algoritme. Sett $\tau_s = 2$ [ms], $p = 0.01$ og $N = \infty$.
- b) La nå antall brukere være endelig, $N < \infty$. Vis at sannsynligheten for å finne filmen blant de N brukerne er $P(X < N) = 1 - (1 - p)^N$.

Videre, det blir nå oppgitt at forventet antall søkesteg gitt at filmen finnes blant de N brukere er

$$E(X|X < N) = E(X) - N \frac{1 - P(X < N)}{P(X < N)}, \quad (1)$$

hvor $E(X)$ er forventet antall søkesteg med $N = \infty$.

- c) Beregn forventet søketid gitt at filmen finnes i P2P nettet med $N = 200$ brukere. Hva blir forventet søketid etter film som finnes i P2P eller i mediedatabasen? Tiden for søk etter film i mediedatabasene, målt etter at et mislykket søk i P2P nettet er gjennomført, er $\tau_{db} = 150$ [ms].

La oss nå anta at en film som er lastet ned og er tilgjengelig i P2P nettverket kan bli utilgjengelig når filmens tidsbegrensede lisens løper ut. Hvis en bruker forespør en film som ikke er tilgjengelig i P2P-nettet vil den sende en forespørsel til mediedatabasen og laste den ned over gatewayen.

Vi gjør følgende antakelser:

- Det er et endelig antall, N_F , filmer tilgjengelig total (i databasene hvorav noen er lastet ned i P2P nettet).
 - Sannsynligheten for å finne en film på P2P nettet øker lineært med antall filmer som er tilgjengelig der, dvs. $p_i = i/N_F$. (i er lik antall filmer i P2P nettet i øyeblikket).
 - Forespørsler etter en film følger en Poissonprosess, med intensitet λ per film.
 - Tiden en film er tilgjengelig i P2P nettet er negativt eksponensialfordelt med intensitet μ
 - P2P nettet kan lagre maksimalt N_P filmer hvor $N_P < N_F$.
- d) Beskriv en Markovmodell av systemet hvor tilstand i er lik antall filmer i P2P nettet. Marker tydelig ankomsts- og avgangsintensiteter og hvordan disse forandres som funksjon av tilstand i . Dette er en kjent modell. Hva kalles denne?
- e) Når $i = N_P$ vil nye forsøk på nedlasting av filmer fra mediedatabasene bli blokkert. Finn et uttrykk for sannsynligheten for at en film ikke kan lastes ned fordi det ikke er plass til flere. Finn et uttrykk for den andelen av tiden nettet er fullt.

Oppgave 2 Pålitelighet til en P2P-tjeneste

I denne oppgaven betraktes tjenester som leveres i P2P domenet vist i figur 1. Når en PC (node i P2P nettet) kan “streame” musikk eller film til de øvrige i nettet regnes den for å være arbeidende (oppe). Hvis ikke regnes den for å ha feilet (nede). At en PC er nede kaller vi her feilet, og at den går fra oppe til nede kalles å feile. Den asymptotiske tilgjengeligheten for en PC er A_{PC} . Feilraten for en arbeidende PC antas og være konstant og betegnes λ_{PC} . Videre antas at alle feil inntreffer uavhengig av hverandre. Det er den enkelte PC-eier som avhjelper feil, og feilavhjelpingen (f.eks. “boote”) kan antas å være uavhengig av tilstanden til de andre PCene i domenet.

En student, tilknyttet P2P nettet via sin PC, planlegger en fest. På festen planlegger hun å vise (“streame”) en film som hun vet ligger lageret på to (2) PCer i P2P domenet og “streame” en “musikk-pakke” som ligger lageret på tre (3) PCer. Anta til å begynne med at PCene hun kan hente film og musikk fra er forskjellige.

Transportdelen av P2P nettet (flytting av informasjon mellom noder/PCer) antas feilfritt.

- Idet festen skal starte, ventes P2P nettet å være i en pålitelighetsmessig stasjonærtilstand mhp. antall arbeidende og feilte PCer. Forklar hva som menes dette. Hva er den pålitelighets-termen for sannsynligheten for at hun får levert film og musikk som ønsket idet festen starter?
- Hun ønsker å bestemme sannsynligheten, U_1 for at musikk eller film ikke leveres ved starten av festen. Tegn et pålitelighetsblokkskjema av systemet som leverer denne tjenesten og finn U_1 uttrykket ved A_{PC} .
- Angi hvilke elementer (PCer) i systemet som inngår i de minimale kuttsettene. Hvis vi antar at $(1 - A_{PC}) \ll 1$, forklar hvordan vi kan finne et rimelig anslag på svaret i punkt b ved å betrakte kuttsettene.

Umiddelbart før festen finner hun at to (2) av PCene som lever musikk og begge (2) PCene som kan levere film er arbeidende (fungerer). Både musikk og film må leveres hele tiden. Festen starter lørdag kl 1800 og arrangøren regner ikke med at noen PCer som er feilet eller feiler vil bli satt i drift igjen mens festen pågår. Den ventes å vare $t_p = 8$ timer.

Du skal nå anta at den ene av PCene som lagrer filmen og den ene av PCene som lagrer musikken fysisk er den samme PCen.

- Finn et uttrykk for sannsynligheten for at festen kan gjennomføres uten at leveransen av film eller musikk feiler. Påpek hvilke av de ovenstående antakelsene du baserer fremgangsmåten på.

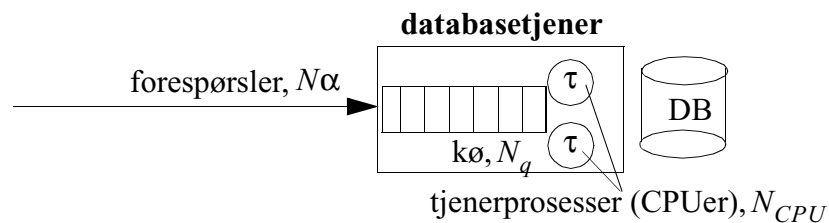
De tre PCene hun henter sin underholdning fra, benyttes til et interaktivt spill. Når en av de tre PCene feiler medfører dette at feilraten på de to resternede øker til $\lambda_G \gg \lambda_{PC}$. Denne raten vedvarer inntil feil i alle de tre PCene er avhjulpet.

Anta i det etterfølgende at feilte PCer blir reparert. Samtlige PCer har samme midlere feilavhjelpingstid.

- e) Forklar kort hvorfor blokkskjema ikke kan benyttes til en pålitelighetsanalyse av systemet under disse forutsetningene. Hvilke(n) antakelse(r) må du gjøre ut over det som er gitt tidligere i oppgaven for at systemet skal kunne modelleres vha. et tilstandsdiagram (en Markov modell)?
- f) Tegn et tilstandsdiagram (en Markov modell) av systemet når tilleggsopplysningene tas i betraktning. (Det er ikke nødvendig å ekspandere modellen etter en systemfeiltilstand er nådd.) Angi tydelig hvilke PCer som er arbeidende og feilte i de ulike tilstandene i systemet samt feil og reparasjonsrater.

Oppgave 3 Simulering av betalingsstrategier

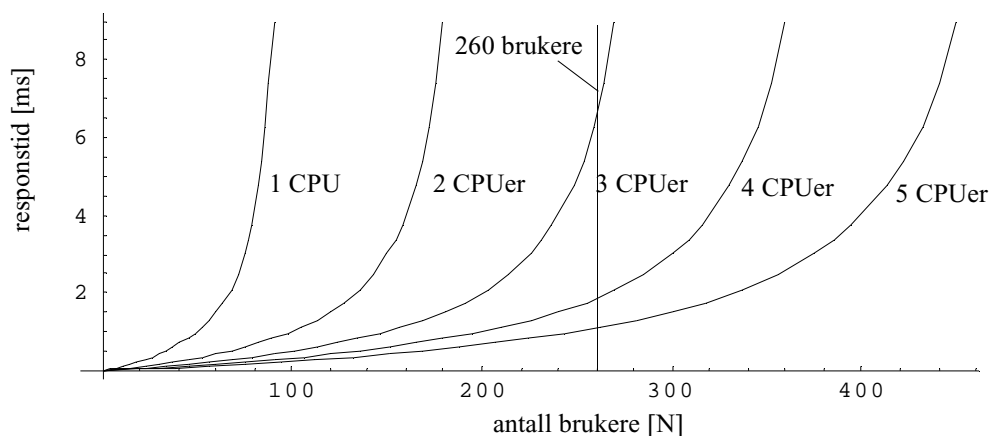
Mediedatabasen har en eller flere tjenerprosesser som behandler innkommende forespørsler, finner fram angjeldende film eller musikk, registrerer og debiterer bruker, koder innholdet og sender det til brukeren. Tjenerprosessene utføres i parallell på hver sin CPU. Anta at den totale tiden en tjenerprosess bruker på å behandle en forespørsel er τ . Forespørsler som ankommer når alle tjenerprosessene er i bruk legges i en kø. Køen er felles for alle forespørsler og tjenesteprosesser. Hvis køen er full, avvises forespørselen. I figur 3 er det vist en skisse av databasetjeneren.



Figur 3 Prinsippskisse av en databasetjener

- a) Beskriv en simuleringsmodell ved hjelp av aktivitetsdiagram av en databasetjener hvor det er mulig å vurdere effekten av å endre antall brukere, N , køstørrelse, N_q , og antall tjenerprosesser (CPUer), N_{CPU} . (Tenk nøye gjennom hvilke detaljer som er tilstrekkelig og nødvendig for å vurdere tjenerens kvalitet under ulike konfigurasjoner). Hvordan kan du endre antall brukere, køstørrelsen, og antall CPUer i modellen din?

Kurveskaren i figur 4 viser kurver basert på resultatene fra simuleringene våre. Vi har her antatt uendelig kø. De viser responstiden som funksjon av antall brukere, en kurve per tjenerprosess(CPU)-konfigurasjon som er testet ut.



Figur 4 Responstid som funksjon av antall brukere for ulike tjenerprosess-konfigurasjoner.

- b) Anta at det er $N = 260$ brukere (markert som en strek i figuren). Spesifiser et akseptabelt øvre responstidsnivå. Hvor mange tjenerprosesser er nødvendig for å oppfylle kravene? Gi en kvalitativ vurdering koplet til valg av responstidsnivå og antall brukere gitt en databasetjener som er konfigurert med alt fra 1 til 5 CPUer når det er $N = 260$ brukere. Forklar hvordan du kan innhente statistik-

ken som danner grunnlaget for kurvene i figur 4. Beskriv dette enten ved referanse til DEMOS primitiver og rapporter, eller markert i aktivitetsdiagrammene dine fra oppgave a).

```

CLOCK TIME =      0.000
*****
*
*           T R A C I N G   C O M M E N C E S
*
*****

TIME/ CURRENT      AND ITS ACTION(S)
0.000 DEMOS        HOLDS FOR 10.000, UNTIL 10.000
                   B1 1      SEIZES 1 OF T
                           HOLDS FOR 2.000, UNTIL 2.000
                   B2 1      AWAITS 1 OF T
                   B3 1      AWAITS 1 OF T
                   B4 1      AWAITS 1 OF T
2.000 B1 1         RELEASES 1 TO T
                           ***TERMINATES
                   B2 1      SEIZES 1 OF T
                           HOLDS FOR 2.000, UNTIL 4.000
4.000              RELEASES 1 TO T
                           ***TERMINATES
                   B3 1      SEIZES 1 OF T
                           HOLDS FOR 2.000, UNTIL 6.000
6.000              RELEASES 1 TO T
                           ***TERMINATES
                   B4 1      SEIZES 1 OF T
                           HOLDS FOR 2.000, UNTIL 8.000
8.000              RELEASES 1 TO T
                           ***TERMINATES

                           CLOCK TIME =      10.000
*****
*
*           R E P O R T
*
*****

                           R E S O U R C E S
                           *****

TITLE      /   (RE)SET/   OBS/ LIM/ MIN/ NOW/   % USAGE/ AV. WAIT/QMAX
T          0.000      4    1    0    1          3

```

Figur 5 Traceresultat fra simulering

I figur 5 er det gjengitt resultater fra en simulering. Resultatene viser tracedetaljer fra en simulering av et system der 4 entiteter (B1-B4) konkurrerer om en felles ressurs (T) med kapasitet lik 1. Hver entitet etterspør én ressurs, bruker denne i T_B tidsenheter (verdien på T_B kan avleses fra tracet), og frigir ressursen og forlater systemet. Kjøringen varte i $T = 10$ tidsenheter.

- c) Lag et tidsplott som viser antallet entiteter i systemet som funksjon av tiden i perioden $(0, T)$. Beregn de to verdiene i rapporten som mangler, % USAGE og AV. WAIT.

Tilgang på filmer og musikk koster penger. Vi skal se på prisprofiler for filmdistribusjon. Ved nedlasting fra en mediedatabase betaler man innholdsleverandøren. Brukerne kan velge mellom 3 mulige opsjoner:

- 40 kroner for ett døgns leie
- 100 kroner for en ukes leie
- 200 kroner for kjøp

Hvis filmen du ønsker å se allerede er lastet ned av noen andre i P2P nettverket så kan du velge å “streame” fra denne brukeren. Det koster 5 kroner/film som betales direkte til den brukeren som har lagret (eier/leier) filmen. Dette betyr at den som laster ned filmen tar en økonomisk risiko, men samtidig kan han tjene penger på videredistribusjonen innenfor P2P nettverket.

Anta at det finnes en søkefunksjon, $søk(film)$, som returnerer en referanse til brukeren hvor filmen finnes, og NULL hvis filmen ikke er funnet. Når en film er lastet ned, lagres informasjon med funksjonen, $lagre(film, CURRENT)$. Filmen vil automatisk forsvinne ved leieperiodens utløp. Informasjon om hvilken film og hvilken bruker blir lagret i et register som funksjonen $søk(film)$ opererer på. Du kan derfor se bort fra håndtering av søk, lagring, registrering og sletting av filmer i din modell.

- d) Lag en simuleringsmodell hvor det er mulig å sammenlikne kostnadene ved valg av ulike strategier. Legg opp til at modellen studerer en spesifikk strategi ad gangen, og at modellen enkelt kan endres slik at andre strategier kan studeres (leie eller eie filmer? vente og “streame” filmer?). En mekanisme for å håndtere flyten av elektroniske penger (betaling for film, inntekter fra videredistribusjon, fylle opp konto) må eksplisitt modelleres. Beskriv modellen ved hjelp av aktivitetsdiagrammer. Mål for sammenlikning kan f.eks. være hva det kostet å laste ned et bestemt utvalg av filmer med bruk av ulike strategier.