



Faglig kontakt under eksamen:  
Poul Heegaard (73 594321 / 99286858)

EKSAMEN I EMNE  
TTM4110 PÅLITELIGHET OG YTELSE MED SIMULERING

Onsdag 6. desember 2006  
Tid: 09:00 – 13:00

Hjelpemidler:

C - Graham Birtwistle: DEMOS - A system for Discrete Event Modelling on Simula. Formelsamling i fag TTM4110 Pålitelighet og ytelse med simulering. NB! Formelsamlingen er vedlagt.

Sensur uke 1, 2007

### Kontekst- og lokasjonsspesifikke tjenester

På NEST airport tilbys et sett av tjenester som kan skreddersys til brukernes behov og preferanser, terminaltype (mobiltelefon, PDA, laptop) og nåværende lokasjon (geografisk posisjon). Eksempler på tjenester er “Show me on airport map”, “mPay at kiosk, shop and cafe”, “Route me to nearest ATM”, og “Order taxi and guide me to taxi stand”. Tjenestene settes sammen av en ASC (Airport Service Composer) som bruker *tjenestekomponentene* LS (Location Server providing the geographical position), CS (Context Server providing the user profile and mobile device characteristics), A4C (check users identity, issues certificates, handles payment), og N (network for transport of information messages). *Tjenestearkitekturen* er gitt i Figur 1.

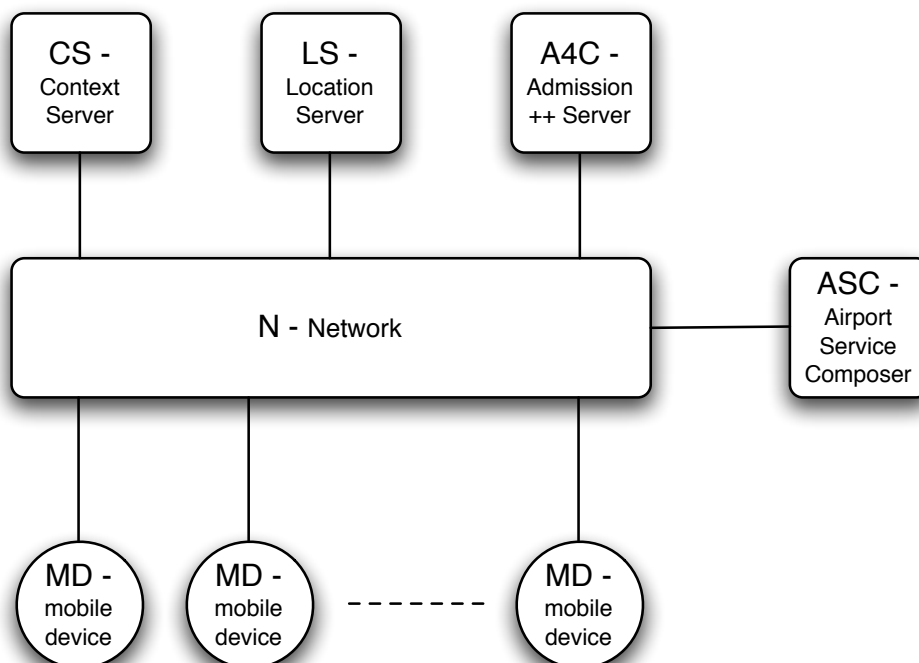


Figure 1: Tjenestearkitekturen

Sekvensen av meldinger for tjeneste *A* (“Show me on airport map”) er vist i Figur 2. Tjenesten bruker fire av tjenestekomponentene beskrevet i tjenestearkitekturen i Figur 1. Sekvensdiagrammet i Figur 2 viser tidssekvensen av meldinger som sendes til og fra de ulike tjenestekomponentene (firkanter på toppen) og tidene  $T_i$  i hver komponent  $i$ . Alle tidene  $T_i$  er negativ eksponentialfordelte med intensitet  $\lambda_i$ ,  $i = MD, N, ASC, LS$ . Den totale tjenestetiden  $T_S$  er tiden fra en forespørsel (request) er sent fra MD til et svar (reply) er mottatt og prosessert i MD.

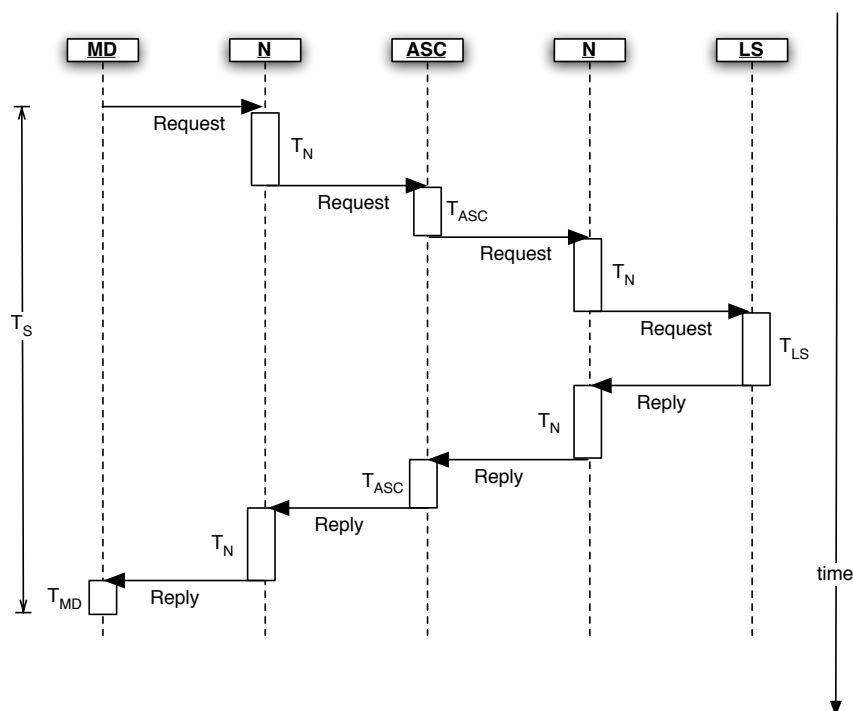


Figure 2: Medlingssekvens for tjeneste A.

- a) Hva er forventet tid  $E(T_{LS})$  i lokasjonstjeneren (LS)? Hva er forventet total tjenestetid  $E(T_S)$  for tjenesten A som beskrevet i Figur 2? La nå  $\lambda_i = \lambda = 0.2 \text{ [ms}^{-1}\text{]}$ , for alle  $i = MD, N, ASC, LS$ . Hva blir sannsynlighetstetthetsfunksjonen til  $T_S$ ? Hva kalles denne funksjonen? Hva er sannsynligheten for at total tjenestetid er mindre enn 50 ms,  $P(T_S \leq 50)$ ?

Answer:

Expected message service time in LS:  $E(T_{LS}) = 1/\lambda_{LS}$

Expected total service time:  $E(T_S) = 1/\lambda_{MD} + 4/\lambda_N + 2/\lambda_{ASC} + 1/\lambda_{LS}$

Probability density function is an Erlang-8 distribution:  $f(t) = \frac{(\lambda t)^7}{7!} \lambda e^{-\lambda t}$ ,  $t \geq 0$

Probability of service time less than 50 ms:  $F(50) = P(T_S \leq 50) = 1 - e^{-0.2 \cdot 50} \sum_{j=0}^7 \frac{(0.2 \cdot 50)^j}{j!} = 1 - 0.22 = 0.78 \text{ (0.779779)}$

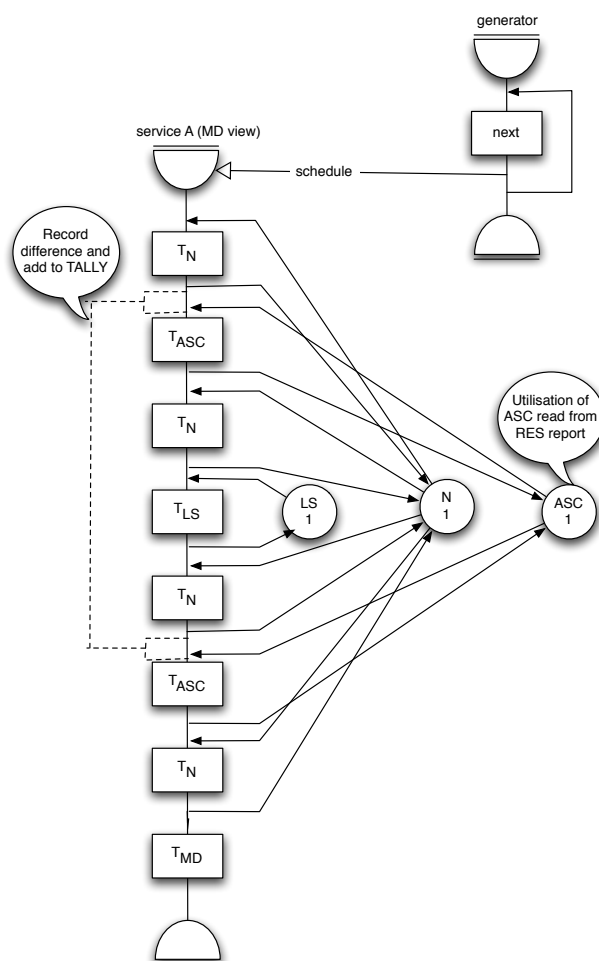
ASC må være en tjener med høy ytelse og høy pålitelighet ettersom den har en sentral rolle i

alle tjenestene som leveres ved flyplassen. Tjeneren skal derfor studeres nærmere med hensyn på både ytelse og pålitelighet. Først skal vi se på ytelsen.

- b) Beskriv ved hjelp av aktivitetsdiagram en simuleringsmodell for å evaluere ytelsen til ASC målt ved tjenerutnyttelsen og kø av forespørsler ved håndtering av tjenesten *A* (se sekvensdiagrammet i Figur 2). Identifiser entiteter og ressurser i modellen. Vis tydelig hvordan du vil samle inn data for å estimere de ikke-funksjonelle egenskapene som skal evalueres.

Answer (the most compact one):

- Entities; the service *A* (as seen from a user) + a generator of service *A* entities
- Resources; Location Server (LS) + Network (N) + Airport Service Composer (ASC) - note that it is not sufficient to have ASC as a resource and forget about LS and N. Both LS and N may have capacity constraints that will influence the service execution times.
- Statistics; it is indicated in the figure below where the start and stop times of the service execution time are recorded. In DEMOS: By use of TALLY you will find the average , standard deviation, max and min values of service execution times in the final report together with the resource utilisation in RES.
- Activity diagram; see figure below



c) I Figur 3 er fire forskjellige sannsynlighetstetthetsfunksjoner plottet. Forklar hvilke av plottene som passer sammen med de følgende tre fordelinger:

1. Punktpøvefordelingen til  $X_i \sim \text{negative exponential distribution}(\lambda = 1)$
2. Fordelingen til gjennomsnittet  $\bar{X}$ , hvor  $\bar{X} = 1/n \sum_{i=1}^n X_i$  med  $n = 40$ .
3. Punktpøvefordelingen til  $N_i \sim \text{geometric distribution}(p = 0.4)$

Answer (norsk versjon, i den engelske er (a) og (b) byttet om):

1 -> (a)

2 -> (b)

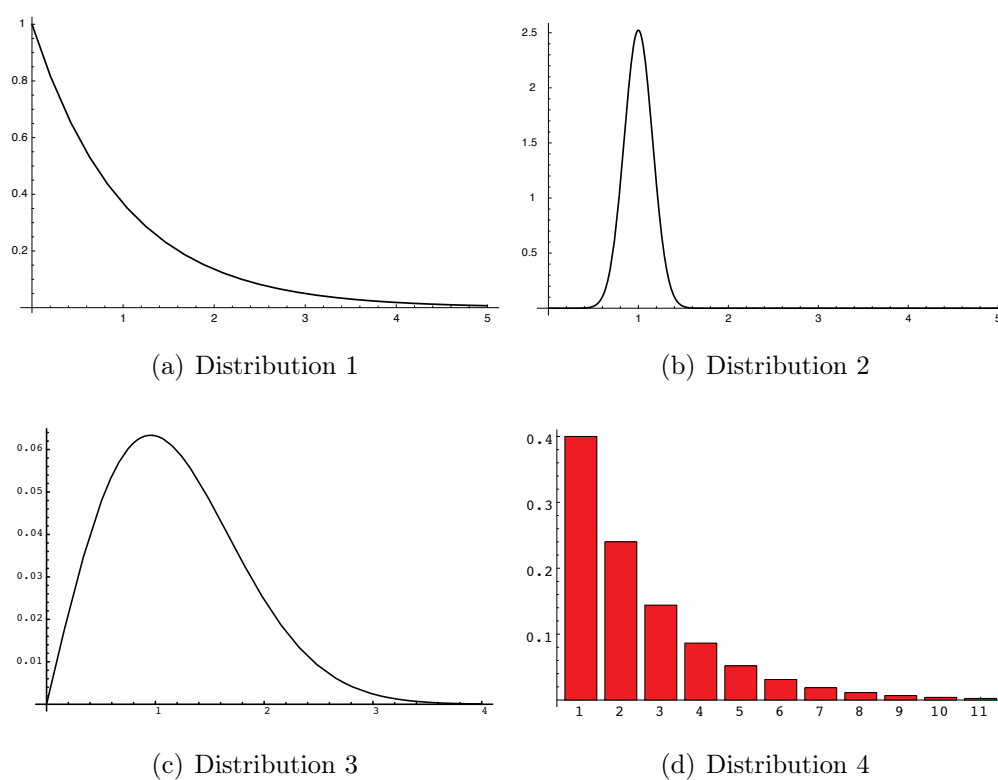


Figure 3: Probability density distributions

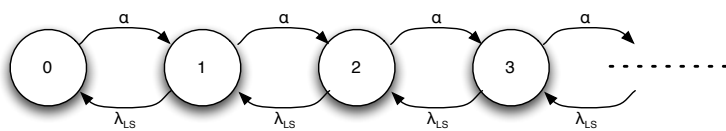
3 -> (d) [it is a *shifted* geometrical distribution]

Lokasjonstjeneren er delt mellom mange bruker, flere tjenester og tjenestetilbydere. Dette betyr at den totale betjeningstiden til LS,  $T_{LS}$ , er avhengig av antall samtidige forespørsler. Anta at kun en forespørsel kan behandles ad gangen og at innkomne forespørsler som møter opptatt tjener legges i en kø med uendelig kapasitet. Forespørsler genereres uavhengig av hverandre med konstant intensitet,  $\alpha$ . Anta FIFO kødisiplin.

- d) Tegn Markov modell av LS køsystem. Bruk Kendalls notasjon for å beskrive hvilken type køsystem dette er. Hva er et stabil kø-system og hva må betjeningsintensiteten være for at kø-systemet i denne oppgaven skal være stabilt? Hva er forventet antall forespørsler i systemet? Hvordan kan systemtiden i LS bestemmes?

Answer:

It is an M/M/1 queue. The service intensity is  $\lambda_{LS}$ . The state diagram is given in:



A queue is stable when the expected number in the system is finite (the queue does not grow to infinity). The condition is that  $A < 1$ , which means that  $\alpha < \lambda_{LS}$ .

Expected number  $E(L) = \sum_{i=1}^{\infty} i p_i = A/(1 - A) = \frac{\alpha}{\lambda_{LS} - \alpha}$  where  $A = \alpha/\lambda_{LS}$

System time is determined by use of Little's result;  $E(L) = E(S)\alpha \Rightarrow E(S) = E(L)/\alpha = \frac{1}{\lambda_{LS} - \alpha}$ .

La oss nå se nærmere på fordelingen av antall forespørsler i LS. Anta stasjonærhet.

- e) Hvis en forespørsel ved ankomst møter  $N$  forespørsler i systemet, hva er sannsynlighetstettheten til system tiden for denne forespørselen? Hva er sannsynligheten for å se  $N$  forespørsler i systemet? Hvordan kan du bestemme systemtiden til en vilkårlig forespørsel?

Answer (Section 6.7.4 on page 167)

(i) Probability density function of system time  $T_N$  given that  $N$  request are in the system on arrival is  $f(t | k) = \text{Erlang-}k$ . The  $T_N$  must wait for  $N + 1$  service completions (including itself).

(ii) Probability of  $N$  requests in an M/M/1 system is  $p_N = (1 - A)A^N$  (score if you know how to obtain this or if you remember this)

(iii) Probability density function of system time  $T$  is  $f(t) = \sum_{k=0}^{\infty} f(t | k)p_k$ .

Lokasjonstjeneren ved NEST airport er kjent for å være ustabil. Sannsynligheten for å få respons fra LS er  $p_{LS}$ . Når LS ikke svarer, venter ASC en konstant tid  $T_D$  før den forsøker igjen.

- f) Hva er sannsynlighetstetthetsfunksjonen til antall forsøk,  $N_{LS}$ ? Hva er forventet antall forsøk,  $E(N_{LS})$ , før en respons er motatt? Hva er forventet responstid  $E(T_S)$  med antagelse om ustabil LS?

Answer:

$N_{LS}$  - number of retries is geometrically distributed

$$f_{N_{LS}}(n) = p_{LS}(1 - p_{LS})^n; n = 0, 1, \dots \Rightarrow$$

$$E(N_{LS}) = \sum_{n=0}^{\infty} n \cdot f_{N_{LS}}(n) = \sum_{n=0}^{\infty} n \cdot p_{LS}(1 - p_{LS})^n = p_{LS} \sum_{n=0}^{\infty} n \cdot (1 - p_{LS})^n = \frac{1 - p_{LS}}{p_{LS}}$$

$E(T_S) = E(T_{MD}) + 2E(T_{ASC}) + 4E(T_N) + E(T_{LS}) + \sum_{n=0}^{N_{max}} f_{N_{LS}}(n) \cdot n \cdot T_D$ , where  $N_{max}$  can be either finite or infinite.

Tjeneste  $B$  er “mPay at kiosk, shop and cafe”. Denne tjenesten avhenger av alle tjenestekomponentene beskrevet i Figur 1 over. Tjenesten er tilgjengelig når alle komponentene er tilgjengelige.

- g) Spesifiser nødvendige betingelser og tegn pålitelighetsblokkskjema (reliability block diagram) for tjenesten  $B$  slik at funksjonssannsynligheten (reliability function),  $R(t)$ , til tjenesten kan bestemmes. Hva er tjenestefeilraten? Bestem, sammenlign og kommentér MTTF, MTFF og MUT for tjenesten  $B$ .

Answer:

Series structure of all components.

Service failure rate (in this case this is the same as the system failure rate);  $\Lambda = \lambda_{MD} + \lambda_N + \lambda_{CS} + \lambda_{LS} + \lambda_{A4C} + \lambda_{ASC}$

Reliability function,  $R(t) = e^{-\Lambda t}$  where  $\Lambda$  is the service failure rate

**MTTF = MTFF = MUT** =  $1/\Lambda$  - they are equal because we have only one upstate



Tjeneste  $C$  ("Route me to nearest ATM") er avhengig av kontekst- (CS) og lokasjonstjenester (LS). Konteksttjenesteren CS på flyplassen returnerer brukerkontekst på 50 [ms]. Hvis CS har feilet så vil ASC rute kontekstforespørsel til en ekstern CS som returnerer svar på 170 [ms]. Lokasjonstjenesteren er avhengig av et stort antall sensorer som er spredt over hele flyplassen. På hvert sted på flyplassen er brukerne innenfor rekkevidden til tre sensorer. La oss nå fokusere på et område ved Gate B. Når alle tre sensorer virker vil lokasjonstjenesteren svare på 10 [ms]. Hvis en sensor har feil er svartiden 40 [ms], og med to feilede sensorer er den 160 [ms]. Hvis ingen sensorer virker er det ikke mulig å få svar, og responstiden er uendelig.

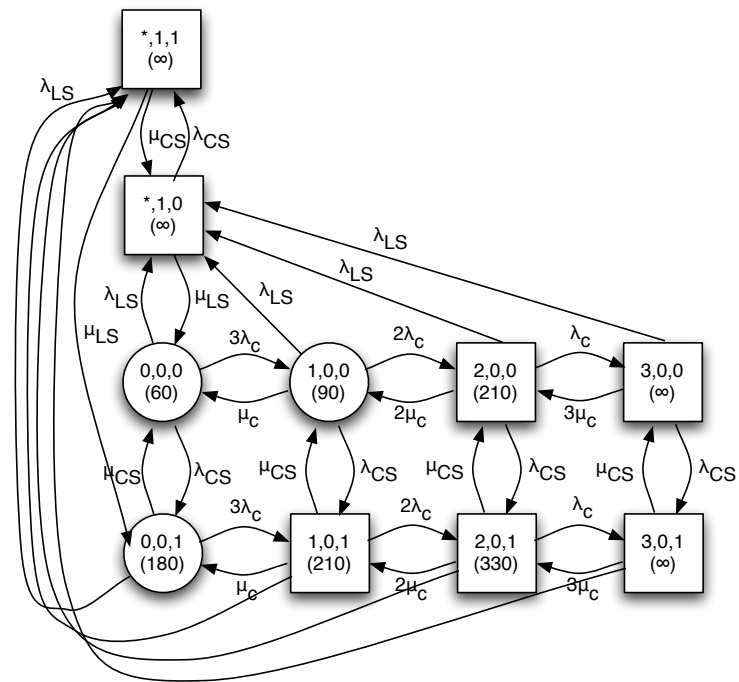
Anta nå at alle svartider fra LS er konstante. Alle komponenter i Figur 1 er perfekte og feiler aldri. Unntaket er CS og LS hvor feil inntreffer uavhengig av tidligere feil og med konstante feilrater, h.h.v.  $\lambda_{CS}$  og  $\lambda_{LS}$ . Feilraten til hver sensor er  $\lambda_c$ . Tilsvarende reparasjonstid er uavhengig, identisk negativ eksponentialfordelt med intensiteter, h.h.v.  $\mu_{CS}$ ,  $\mu_{LS}$  og  $\mu_c$ . Den eksterne CS feiler aldri. Reparasjon av sensorer initieres ved feil, og når en feilet LS repareres så repareres også alle feilede sensorer i samme operasjon.

- h) La tjenesteresponstiden for tjeneste  $C$  være maksimalt  $T_{C,max} = 180$  [ms]. Spesifiser tilstandene til leveranse av tjeneste  $C$  og identifiser nedetilstander for denne tjenesten. Tegn Markov modell for tjenesten hvor tilstandssannsynlighetene og tjenesteutilgjengeligheten kan bestemmes. Vis hvordan du etablerer tilstandsligninger i denne modellen (vis én ligning).

Answer:

State definition; (# of failed sensors, # of failed internal LS, # of failed internal CS)

State diagram: The number of failed sensors does not count when the LS has failed.



State equation; one line

$$p_{1,0,0}(\lambda_{LS} + 2\lambda_c + \mu_c + \lambda_{CS}) = 3\lambda_c p_{0,0,0} + 2\mu_c p_{2,0,0} + \mu_{CS} p_{1,0,1}$$

Remember the normalisation condition,  $\sum_{i=0}^3 \sum_{j=0}^1 \sum_{k=0}^1 p_{i,j,k} = 1$ .