

NTNU
Norges teknisk-naturvitenskapelige
universitet

Fakultet for informasjonsteknologi,
matematikk og elektroteknikk

Institutt for datateknikk
og informasjonsvitenskap



Kontinuasjonsløsning på SIF8037 Distribuerte systemer og ytelsesvurdering
(Distribuerte systemer kun)

Fredag 16. august 2002, 0900-1400

Der det synes å mangle noen opplysninger, må det angis hvilke antagelser som synes å være naturlige.

Distribuerte systemer dekkes av oppgavene 1, 2, 3 og 4, mens ytelsesvurdering dekkes av oppgavene 5, 6, 7 og 8.

Oppgave 1 – Kommunikasjon og synkronisering (Communication and synchronization) – 12,5%

- a) Angi kort hvilke hovedoppgaver som må løses i forbindelse med kommunikasjon og synkronisering i et distribuert system

SVAR:

Kommunikasjon & Synkronisering - nytt i distribuerte systemer kontra sentraliserte systemer:
Felles tidsbegrep

Kommunikasjon - nytt i distribuerte systemer kontra sentraliserte systemer:
RPC / RMI – mekanismer (m/ tilhørende parameter-, binding-, kontakt- og feilhåndtering)
Gruppemekanismer (m/tilhørende atomiskhet og ordnethet)

Synkronisering - nytt i distribuerte systemer kontra sentraliserte systemer:
Gjensidig utelukkelse - med nye krav
(unngå enkeltressursavhegighet, unngå/begrens flaskehalser, begrensn meldingsmengder)
Primasutvelgelse - med ny funksjonalitet
(velg en bestemt sekretær fra en gruppe)

- b) Beskriv konsist hvordan logiske klokke (logical clocks) brukes til å løse noen av disse oppgavene

SVAR:

Logiske klokke anvender N endimensjonale tellere
Tidsmerker T_i / T_j brukes til synkronisering av slike endimensjonale tellere $L_i; i: 1 \dots N$

Hendelse – På node I (inkl. / ekskl. motta/send) $L_i := L_i + 1$
Meldingssending – Fra node I til node J: $T_j := L_i$
Meldingsmottak – Til node I fra node J: $L_i := \text{MAX} (L_i , T_j)$

Kommunikasjon: Logiske klokke brukes her for passiv ordning av hendelser
Synkronisering: Alle typer (sentraliserte/distribuerte/ringbaserte)
anvender logiske klokke for aktiv ordning av hendelser

- c) Identifiser minst to andre mekanismer som anvendes til å løse slike oppgaver

SVAR:

Fysiske klokke
Vektorklokke

- d) Sammenlign fordeler og ulemper med hver av de mekanismene som du har diskutert i deloppgave b) og c)

SVAR:

- Fysiske klokker: Per definisjon unødvendig & umulig å implementere korrekt, svært ressurskrevende overhead
- Logiske klokker: Holder styr på hendte-før relasjoner & antall hendelser for alle samlet, lite ressurskrevende overhead
- Vektorklokker: Holder styr på hendte-før relasjoner & antall hendelser for hver enkelt, mer ressurskrevende overhead

Oppgave 2 – Distribuerte filsystemer (Distributed file systems) – 12,5%

- a) Angi kort hvilke hovedoppgaver som må løses i forbindelse med den fysiske avstand mellom hvor en fil lagres og hvor den kan aksesseres i et distribuert filsystem

SVAR:

Implementere rask og trygg nok aksess til dataene

- b) Beskriv konsist hvordan caching brukes til å løse noen av disse oppgavene

SVAR:

- Ved uthenting: Lage en ny eller oppdatere en eksisterende temporær kopi av dataene lokalt
- Opp oppdatering: Lagre endringene i den lokale temporære kopien, og overføre fra lokal kopi til global forekomst ved behov eller med jevne mellomrom

- c) Identifiser minst en annen mekanisme som anvendes til å løse slike oppgaver

SVAR:

Replisering

- d) Sammenlign fordeler og ulemper med hver av de mekanismene som du har diskutert i deloppgave b) og c)

SVAR:

Caching

Fordeler: Gir økt ytelse

Ulemper: Kan gi konsistensproblemer – med for mange versjoner og for ofte endringer

Replisering

Fordeler: Gir økt pålitelighet

Ulemper: Kan gi minnet ytelse – med for mange kopier og for mange oppdateringer

Oppgave 3 – Distribuerte databasesystemer (Distributed database systems) – 12,5%

- a) Angi kort noen deloppgaver som må løses i forbindelse med håndtering av feil og autonomi i et distribuert databasesystem

SVAR:

Feil: Må håndtere at ulike noder fungerer, mens andre noder feiler (uten å stoppe systemet)

Autonomi: Må akseptere at noder tar avgjørelser over forhold som har med lokale formål å gjøre

- b) Beskriv konsist hvordan tofase bekrefting (2PC – 2 phase commit) brukes til å løse noen av disse oppgavene

SVAR:

Todelt avslutning: 1.fase m/stemming & 2 fase m/utføring –
etter alle-har-veto prinsippet koplet til enten feil eller autonomi

- c) Identifiser minst to andre algoritmer som anvendes til å løse slike oppgaver

SVAR:

2PW - 2fase skrivning

3PC - 3fase bekrefting

- d) Sammenlign fordeler og ulemper med hver av de algoritmene som du har diskutert i deloppgave b) og c)

SVAR:

2PW: Håndterer feil innen en node på en effektiv måte, håndterer ikke feil mellom noder

2PC: Sikrer autonomi og effektivitet i normaltilfeller uten mye feil, mindre overhead

3PC: Sikrer autonomi og effektivitet i unntakstilfeller med mye feil, mer overhead

Oppgave 4 – Distribuerte navnetjenester (Distributed name services) – 12,5%

- a) Angi kort noen deloppgaver som må løses i forbindelse med navigering mot navn i en distribuert navnetjeneste

SVAR:

Må kunne håndtere navnerom med uendelig mange navn som er i kontinuerlig endring

- b) Beskriv konsist hvordan iterativ navigering brukes til å løse noen av disse oppgavene

SVAR:

Klienten selv mapper del for del av navnet i en sekvensiell prosess mot ulike tjenere

- c) Identifiser minst en annen algoritme som anvendes til å løse slike oppgaver

SVAR:

Rekursiv navigering

- d) Sammenlign fordeler og ulemper med hver av de algoritmene som du har diskutert i deloppgave b) og c)

SVAR:

Iterativ navigering:

Mer klientbelastende, kan lettere utnytte caching, autorisering blir vanskeligere

Rekursiv navigering:

Mer tjenerbelastende, kan lettere implementere autorisering, caching blir vanskeligere