

**Norges teknisk-naturvitenskapelige universitet  
Institutt for datateknikk og informasjonsvitenskap**



**EKSAMENSOPPGAVE I FAG TDT4150 – AVANSERTE DATABASESYSTEMER**

**Faglig kontakt under eksamen: Kjetil Nørvåg**

**Tlf.: 93440**

**Eksamensdato: 7. juni 2010**

**Eksamenstid: 09.00-13.00**

**Tillatte hjelpemiddel: D: Ingen trykte eller håndskrivne hjelpemiddel tillatt. Bestemt, enkel kalkulator tillatt.**

**Språkform: Bokmål**

**Sensurdato: 28. juni 2010**

### Oppgave 1 – Parallele og distribuerte databaser – 20 %

- Hva er hhv. *speedup* og *scaleup*?
- Vil det alltid være fornuftig å dele en relasjon opp i like store partisjoner?  
Hvorfor / Hvorfor ikke?
- Har valget mellom synkron og asynkron oppdatering av replikater noen konsekvenser for sannsynligheten for vranglås? Begrunn svaret.

### Oppgave 2 – Peer-til-peer – 20 %

Beskriv Pastry (få med ruting-tabellenes innhold og struktur, hvordan noder kan bli del av et Pastry-nettverk, og hvordan data blir satt inn og gjenfunnet).

### Oppgave 3 – Datavarehus og datagruvedrift – 30 %

- Diskuter forskjeller mellom OLTP (*on-line transaction processing*) og OLAP (*on-line analytical processing*) som gjør at man ønsker separate systemer for datavarehus.
- Anta handlekurv-data til høyre. Bruk apriori-algoritmen for å finne hvilke assosiasjonsregler som gjelder, gitt at minimum støtte er 22 % (dvs. *minimum support count* er 2) og konfidens er 70 %.

TransaksjonsID	Kjøpte produkt
T100	A,B,E
T200	B,D
T300	B,C
T400	A,B,D
T500	A,C
T600	B,C
T700	A,C
T800	A,B,C,E
T900	A,B,C

### Oppgave 4 – Diverse – 30 %

- Er det likegyldig hvilken relasjon som er ytre og indre relasjon ved utføring av join?  
Hvorfor / Hvorfor ikke?
- Er kolonnebaserte databasesystemer slik som C-Store optimalisert for lesing eller skriving?  
Begrunn svaret.
- Anta at man i et databasesystem med støtte for top-*k* join-spørringer ønsker at disse spørringene enkelt skal kunne uttrykkes i SQL. Gi et eksempel på en mulig måte å kunne uttrykke en slik spørring i SQL.
- Forklar hva vi mener når vi sier at et punkt er dominert av et annet punkt.