

Kirjoita jokaisen palauttamasi paperin ylälaitaan selvästi: "T-106.5240, 16.12.2008", koko nimesi, opiskelijanumerosi ja koulutusohjelmiasi sekä palauttamiesi **paperien kokonaismäärä**.

1. (6p) Selitä lyhyesti seuraavat käsitteet:
 - a) pääkopiotoisintaminen (primary-copy replication, master-slave replication)
 - b) toisinteiden heikko keskinäinen johdonmukaisuus (weak mutual consistency of replicas)
 - c) kyselyn sisäinen rinnakkaisuus (intraquery parallelism)
 - d) sivupalvelin (page server)
 - e) jonotettu transaktionkäsittely (queued transaction processing)
 - f) transaktiomonitori (transaction-processing monitor)
2. a) (4p) Kuvaa kaksivaiheinen sitoutumiskäytäntö (two-phase commit protocol) kurssimonisteessa esitetyllä tavalla (eli keskeytysoletukseen, presumed abort, perustuen). Mitä viestejä eri pisteiden välillä liikkuu? Mitä lokikirjauksia tehdään? Milloin loki pakotetaan levyille? Oleta, että transaktio sitoutuu onnistuneesti eikä mikään piste romahda.
b) (2p) Missä tilanteessa kaksivaiheisen sitoutumiskäytännön osallinen (cohort, participant) voi joutua odottamaan pitkäänkin ennen kuin pystyy joko sitouttamaan tai keskeyttämään oman alitransaktionsa?
3. Tarkastellaan päätösvaltakäytäntöön (quorum consensus protocol) perustuvaa toisintetun tietokannan hallintaa, kun kullakin tietoalkiolla on n toisinnetta, lukuvalian koko on p (read quorum) ja kirjoitusvallan koko on q (write quorum).
 - a) (2p) Mitkä ehdot $n:n$, $p:n$ ja $q:n$ tulee täyttää, jotta päätösvaltakäytännön soveltaminen on mahdollista?
 - b) (1p) Valitse p :lle ja q :lle arvot, kun yhden pisteen romahdus ei saa estää käytännön soveltamista ja kun $p:n$ tulisi olla mahdollisimman pieni.
 - c) (1p) Oletetaan, että $n = 20$. Mikä on $p:n$ ja $q:n$ pienin mahdollinen arvo silloin kun $p = q$?
 - d) (2p) Oletetaan, että $n = 20$. Valitse p :lle ja q :lle arvot, kun suurin mahdollinen määrä pisteitä voi romahtaa käytännön soveltamisen estymättä. Mikä on tämä suurin mahdollinen määrä?
4. a) (4p) Anna tilanne ja SQL-kysely, jonka laskemisessa kannattaa käyttää puoliliitosoptimointia (optimizing with semijoins, planning with semijoins). Kerro lyhyesti, miten kyselysi lasketaan ja missä kohdassa tehdään puoliliitos.
b) (2p) Miten toimii epäsymmetrinen paloittelu- ja toisinnusliitos (asymmetric fragment-and-replicate join)?
5. Miten seuraavat kyselyt kannattaa rinnakkaistaa rinnakkaistietokannassa? Attribuutin A arvot ovat jakautuneet tasaisesti välille $[0..1000]$. Relaatiot $r(\underline{A}, B)$ ja $s(\underline{A}, C)$ on osaväliositettu (range-partitioned) samoilla osaväleillä kaikille prosessoreille. Kyselyn tulos halutaan yhdelle prosessorille.
 - a) (1p) `select count(*) from r.`
 - b) (1p) `select count(distinct A) from r.`
 - c) (2p) `select distinct B from r where A > 700.`
 - d) (2p) `select distinct C from r, s where r.A = s.A and B < 20.`