Tietokannan hallinta, kevät 2002 **Harjoitus 3 (15.-19.4.)**

Läsnäolo: 3 tehtävää **(**)-merkitty tehtävä lasketaan kahdeksi**

1. (**) Olkoon EMPLOYEE-relaatiossa 10000 monikkoa, monikon pituus 100 merkkiä ja rakenne kuten luennoissa ja kirjassa on esitetty. Relaatio on toteutettu kasana, jaksonpituus on 4 KB. Relaatiolle on tehty kaksi ISAM-rakenteista oheishakemistoa, toinen SSN-attribuutille (pituus 9 B) ja toinen ADDRESS-attribuutille (pituus 40 B). Hakemistojen lehtitasolla (jota luennoilla tuli sanotuksi 'pohjatasoksi' (s. 32)) käytetään täyttöastetta 80%.
a) Laske relaation ja hakemistojen tilantarve jaksoina (samaan tyyliin kuin luentojen sivulla 28 eli lukuarvoja voi laskennan helpottamiseksi pyöristellä, mutta koko hakemiston rakenne jakson tarkkuudella pitää selvittää tarkasti).
b) Missä suhteessa alle 100% täyttöaste on hyödyllinen? Pitäisikö kasankin toteutuksessa käyttää alempaa täyttöastetta kuin 100%?
c) Onko ADDRESS-hakemisto järkevä kyselyjen kannalta (yritystietokannan esimerkkiosoitteet, suomalaiset osoitteet jne) ?

2. Tehdään EMPLOYEE-relaatiolle (oletukset kuten tehtävässä 1) hajautusrakenteinen oheishakemisto attribuutille DNO. Esitä hakemiston rakenne ja laske sen tilantarve (mahdolliset lisäoletukset voit tehdä itse).

3. Toteutetaan relaatio WORKS_ON kasana ja rakennetaan sille hajautusrakenteinen oheishakemisto avaimena ESSN.
a) Vertaa tätä ratkaisua viikon 2 tehtävässä 4 esitettyyn hajautusrakenteisen tiedoston toteutukseen (parempi / huonompi, missä suhteessa?)
b) Kuinka relaatio pitäisi toteuttaa, jos tarvitaan ryhmittelyä sekä projekteittain että henkilöittäin?

4. Oletetaan, että relaatiolle EMPLOYEE on olemassa tehtävissä 1 ja 2 mainitut hakemistot. Mitä levyhakuja (tiedostoon tai hakemistoihin) sisältyy seuraaviin operaatioihin?

insert into EMPLOYEE values
('Jim','B','Koch', 223344556, '10-JAN-50', '635 Voss, Houston, TX', 'M', 35000, 123456789, 4);
delete from EMPLOYEE where SSN = '123456789' OR SSN = '333445555';
update EMPLOYEE set SALARY = SALARY * 1.05 where SSN = '123456789';

5. (**) Oletetaan, että kertaluokan B+ -puuhun (kaikissa solmuissa 2..4 avainta) lisätään järjestyksessä avainarvot 23, 65, 37, 60, 46, 92, 48, 71, 56, 59, 18, 21, 10, 74, 78, 15, 16, 20 ja 24.
a) Esitä rakenteen muodostuminen vaihe vaiheelta.
b) Kuinka B+ -puu muuttuu, kun siitä poistetaan avaimet 24, 23, 10 ja 20?
(Huom. B+ -puun lisäys- ja poisto-operaatiot vastaavat pitkälle luentomateriaalissa esitettyjä B-puun vastaavia operaatioita. B+ -puussa sisäsolmujen avaimia tarvitaan vain 'viitoiksi' ohjaamaan puussa etenemistä.)

6. Tehtävän 1 mukaiset ADDRESS- ja SSN-hakemistot voidaan toteuttaa myös B+ -puuna.
a) Esitä ADDRESS-hakemiston lehti- ja sisäsolmun rakenne tässä tapauksessa.
b) Mikä on ADDRESS-hakemiston tilantarve?
c) Mitä eroa B+ –puuhakemistolla ja ISAM-hakemistolla on kyselyjen (select.., insert.., delete.., update..) kannalta? (Kyselyjen monipuolisuuden kannalta on luontevampaa tarkastella SSN-hakemistoa kuin ADDRESS-hakemistoa.)