IT350-DZ12

1. **Zadatak - Za bazu podataka koju ste kreirali u lekciji 8. pretpostavite da su neke od tabela distribuirane i to na opisani način.**

**Propusnost je 50 000 bita u sekundi, veličina torke je 100b, dok je kašnjenje 0,1s.**

**BAZA PODATAKA 1:**

**LOKACIJA I1: tabela Pacijent (Ime , prezime, JMBG, adresa, telefon, imeDoktora)**

**LOKACIJA I2: tabela Bolest (Naziv, Opis, Slika)**

**LOKACIJA I3: tabela BolujeOd (Pacijent.Ime, Bolest.Naziv, datum\_dijgnoze))**

**Neka su veličine tabela date sa:**

**Pacijent (Ime , prezime, JMBG, adresa, telefon, imeDoktora) – 1 000 n-torki**

**Bolest (Naziv, Opis, Slika) - 500 n-tork**

**BolujeOd (Pacijent.Ime, Bolest.Naziv, datum\_dijgnoze)) - 10000 n-torki**

**Prateći korake osnovne strategije distribuiranog izvršavanja upita, treba sračunati vreme izvršenja upita kojim će se naći svi pacijenti kojima broj telefona počinje sa 065, opise njihovih bolesti i datum dijagnoze, a pod pretpostavkom da se da se upit izvršava na lokaciji I3. Prenos relacija na druge lokacije vršiti po sopstvenom izboru koji treba opisati u rešenju.**

**1**. Prvi korak osnovne strategije distribuiranog izvršavanja ovog upita: sastoji se u izvršavanju, na lokaciji I1, jednorelacionog upita:

SELECT \* FROM Pacijent WHERE telefon LIKE '065%'; 🡪 pretpostavka da upit vraća 200 torki!

i u izvršavanju, na lokaciji I2, jednorelacionog upita:

SELECT opis FROM Bolest;

Prvi upit proizvodi relaciju P'(\*) na lokaciji I1, a drugi upit proizvodi relaciju B'(Opis) na lokaciji I2.

**2**.Drugi korak osnovne strategije postavlja se sledeći izbor koji je definisan u zadatku a to je preneti relacije P' i B' na lokaciju I3 i tamo izvršiti oba spajanja i projekciju na izlazne attribute.

SELECT P'.Ime, B'.Opis, BolujeOd.datum\_dijagnoze FROM P', B', BolujeOd WHERE P'.Ime=BolujeOd.Ime AND BolujeOd.Naziv=B'.Naziv;

t = 0,1 + (700 \* 100) / 50 000 = 0,1 + 1,4 = 1,5 sekundi

-Druge dve opcije su: 1. preneti relacije P' i BolujeOd na lokaciju I2,

2. odnosno preneti relacije B' i BolujeOd na lokaciju I1

1. t = 0,1 + (10200 \* 100) / 50 000 = 0,1 + 20,4 = 20,5 sekundi
2. t = 0,1 + (10500 \* 100) / 50 000 = 0,1 + 21 = 21,1 sekunda

Iz ovih rezultata zaključujemo da je najoptimalnija strategija za rešavanje ovog distribuiranog upita prenošenje rezultata jednorelacijskih upita na lokacijama I1 i I2 i zatim njihovo prenošenje i spajanje na lokaciji I3.

1. **Zadatak** - Navesti što je potrebno uraditi da bi se JOIN dve tabele efikasno obavio na računaru sa više procesora?

Da bi obavili spajanje (JOIN) Tabela1 (att1, att2) i Tabela2 (att2 , att3 ): potrebno je da heširamo torke tabela Tabela1 i Tabela2 na broj skupova – odnosno grupa koji je jednak broju procesora. Međutim, hešing funkcija h koju koristimo mora da zavisi samo od atributa att2 jer se on i nalazi u obe relacije, ne od svih atributa, tako da torke koje se spajaju se uvek šalju u isti skup-grupu. Kao i kod unije, mi šaljemo torke skupa i u procesor i. Tada možemo izvršiti join u svakom procesoru korišćenjem bilo kog jedno procesorskog algoritma za spajanje.

Tomislav Živadinović 3948