

FIZIČKO PROJEKTOVANJE PODATAKA - Tehnike smeštaja podataka

Polazi se od toga da se pristup podacima (čitanje-pisanje) na masovnoj memoriji (M2) obavlja u blokovima (disk-blok). Pri tome:

- ✓ suviše mali blok dovodi do fragmentacije podataka na disku;
- ✓ suviše veliki blok dovodi do dužeg vremena prenosa podataka i slabog iskorišćenja prostora;
- ✓ u uslovima intenzivnog višekorisničkog rada kod velikih IS gubi se efekat fragmentacije.

Na osnovu prethodnog proizilazi zahtev:

Zahtevani podaci treba da su raspoređeni na što manji broj blokova.

Uporedo se primenjuju dve tehnike smeštaja podataka:

- ✓ **distribucija podataka;**
- ✓ **alokacija prostora.**

Distribucija podataka

Globalna

- na odvojene računarske sisteme u mreži

Fragmentaciona

Replikativna

Lokalna

- unutar jednog računarskog sistema

Fragmentaciona

Replikativna

Fragmentaciona distribucija podrazumeva unikatnu distribuciju – podaci su zapisani samo jednom, unutar određenog bloka

| | |
|-----------------|---|
| Motivi za to su | povećanje performansi povećanje konkurentnosti |
|-----------------|---|

| | |
|---------|---|
| Teškoće | distribuiranost transakcija distribuiranost oporavka |
|---------|---|

Replikativna distribucija podrazumeva višestruku distribuciju – podaci su zapisani duplo ili N puta, unutar N blokova koji se nalaze na različitim lokacijama (različite disk-jedinice lokalno ili različiti računarski sistemi (globalno).

| | |
|-----------------|--|
| Motivi za to su | povećanje performansi povećanje konkurentnosti povećanje pouzdanosti |
|-----------------|--|

| | |
|---------|---|
| Teškoće | distribuiranost transakcija distribuiranost oporavka sinhronizacija replika |
|---------|---|

Ovo sve treba da se realizuje na sloju fizičke organizacije podataka (nije standardizovano). Treba da su zadovoljena tri principa

- ✓ **transparentnost fragmentacije**: korisnik (aplikacija) ne sme da bude svestan toga da li je i kako izvršena fragmentacija;
- ✓ **transparentnost replikacije**: korisnik (aplikacija) ne sme da bude svestan da li je i kako izvršena replikacija;
- ✓ **transparentnost lokacije**: korisnik (aplikacija) ne sme da bude svestan gde se (na kom disku, računarskom sistemu) nalaze podaci.

Postoje tri načina fragmentacione distribucije relacione baze podataka:

- ✓ **totalna fragmentacija**: tabele se nalaze na više lokacija s tim da je svaka tabela na samo jednoj lokaciji;
- ✓ **horizontalna fragmentacija**: celi redovi tabele se nalaze na više lokacija prema nekom kriterijumu pripadnosti;
- ✓ **vertikalna fragmentacija**: cele kolone tabele se nalaze na više lokacija, pri čemu svugde mora da se nalazi i primerni ključ.

Kriterijumi pripadnosti redova tabele fragmentu:

- ✓ **Kružni**: redovi se redom kojim nastaju raspoređuju "u krug" na različite lokacije (disk-jedinice, računarski sistemi);
- ✓ **Hash**: redovi se raspoređuju na različite lokacije primenom funkcije transformacije nekog zadržaja u adresu;
- ✓ **Opseg**: redovi se raspoređuju na različite lokacije prema opsegu vrednosti jedne (retko više) kolone.

| | |
|--------------------|---|
| Alokacija prostora | - unutar jedne DATABAZE (bez alokacije) |
| PARTITION | - particija |
| TABLESPACE | - podparticija |
| CLUSTER | - zajedničko područje ("grozd") |
| TABLE | - tabela u grozdu |
| INDEX | - indeks u grozdu |
| TABLE | - tabela izvan grozda |
| INDEX | - indeks izvan grozda |

TABLE i INDEX mogu biti unutar nekog CLUSTER-a ili izvan njega.

Za sve jedinice alokacije (PARTITION, TABLESPACE, CLUSTER) zadaju se:

- ✓ **inicijalni prostor**: kontinuelan, dodeljuje se prilikom kreiranja;
- ✓ **naknadni prostor**: kontinuelan, dodeljuje se kada ustreba dodatni prostor.

U CLUSTER se smeštaju podaci koji se zajedno koriste:

- ✓ redovi tabela koji zadovoljavaju uslov spajanja;
- ✓ redovi tabela i odgovarajuće vrednosti indeksa.