

FIZIČKO PROJEKTOVANJE PODATAKA - INDEKSI - Osnove

Osnovna namena indeksa jeste da se kod oba vida postupaka nad podacima (održavanje i korišćenje) ubrza pristup podacima. Kao i kod svakog kompromisa, to ima svoju cenu: potreban je dodatni memorijski prostor za dodatne pristupne datoteke.

Vrste indeksa nad tabelama

Postoje dve vrste indeksa:

H-indeks (hash) - zasnovan na transformaciji pristupnih podataka (ključa) u adresu. Pristupna datoteka sadrži primarno i sekundarno područje sa ulančanim slogovima strukture <PristupniPodaci,Sledeci,Adresa>

Ovakav indeks omogućava: direktan pristup redu tabele.

Većina implementacija RBP sama pravi H-indekse po primarnim ključevima.

S-indeks (sort) - zasnovan na traženju pristupnih podataka (ključa) u strukturi tipa binarnog stabla po B+ ili B* organizaciji čije čvorove čine stranice koje sadrže sortirane pristupne slogove strukture <PristupniPodaci,Adresa>.

Ovakav indeks omogućava: direktan pristup redu tabele,
logički redosled (uređenost) tabele.

Vrste traženja kod S-indeksa

hard-seek Traženje po odnosu "jednako" kriterijuma i sadržaja reda. Traži se prvi red u tabeli koji je u odnosu == prema kriterijumu traženja.

```
XXXXXXXXX
MMMMMMMMM      ! - |
HHHHHHHHH      |
GGGGGGGGG      |
GGGGGGGGG <- |   |
CCCCCCCCC      |   |
BBBBBBBBB      |   |
AAAAAAA      |   |
              |   |
              G   L
              ok  X
```

Kod traženja G nalazi se prvi red koji ispunjava kriterijum.

Kod traženja L ne nalazi se ni jedan red koji ispunjava kriterijum. Traženje se prekida nailaskom na prvi red kod koga je kriterijum traženja "veći" od zadatog (za rastući indeks, a za opadajući je obrnuto).

soft-**seek** Traženje po odnosu "jednako ili manje/veće" kriterijuma i sadržaja reda. Traži se prvi red u tabeli koji je u odnosu \geq (za rastući indeks) ili \leq (za opadajući indeks) prema kriterijumu traženja.

XXXXXXXX				Kod traženja G nalazi se prvi red koji
MMMMMMMM		<-		ispunjava kriterijum.
HHHHHHHH				
GGGGGGGG				Kod traženja L nalazi se prvi red koji je veći
GGGGGGGG	<-			u odnosu na kriterijum.
CCCCCCCC				
BBBBBBBB				
AAAAAAAA				
	G	L		
	ok	ok		

Osnovni mehanizmi ubrzavajućeg efekta indeksa

M1 Direktni pristup jednom redu po unikatnom kriterijumu.

XXXXXXXX		Kod traženja G nalazi se ili se ne nalazi red koji
MMMMMMMM		ispunjava kriterijum.
HHHHHHHH		
GGGGGGGG	<-	
CCCCCCCC		
BBBBBBBB		
AAAAAAAA		Mogući indeksi su i hash (brže) i sort.
	G	

M2 Selektivni pristup grupi redova po kriterijumu

XXXXXXXX		Kod traženja G nalazi se ili se ne nalazi prvi red koji
HHHHHHHH		ispunjava kriterijum. Ako je red nađen, pristupa se
HHHHHHHH	x	narednim redovima sve do nailaska na red koji ne
GGGGGGGG		ispunjava kriterijum ili do kraja tabele.
GGGGGGGG	-	
BBBBBBBB		
BBBBBBBB		
AAAAAAAA		Mogući indeks je samo sort.
	G	

M3 željeni redosled redova po kriterijumu uređenosti

XXXXXXXXX	^	Pristupa se svim redovima tabele po željenom redosledu određenom indeksom.
HHHHHHHH		
HHHHHHHH		
GGGGGGGG		
GGGGGGGG		
BBBBBBBB		
BBBBBBBB		
AAAAAAA		Mogući indeks je samo sort.

Napomena: M1 je u RBP automatski obezbeđen za primarni ključ.

Osnovne situacije primene indeksa

S1 Direktan pristup željenom redu tabele po unikatnom kriterijumu traženja K

Nad tabelom treba kreirati indeks po kriterijumu K (najčešće je primarni ključ ili kandidat-ključ). Kriterijum K se javlja u WHERE klauzuli upita.

Implementira: M1

S2 Postojanje/nepostojanje reda po kriterijumu traženja K

Ovo je situacija koja se javlja kod upita koji u WHERE klauzuli sadrže podupit kao argument klauzule EXISTS ili NOT EXISTS.

Ako je podupit nad tabelom S i sadrži kriterijum K u WHERE klauzuli, nad tabelom S treba kreirati indeks po kriterijumu K.

Implementira: M1 ako je kriterijum K unikatan,
 M2 ako je kriterijum K višestruk.

S3 Uređenost rezultata upita po kriterijumu uređenosti K

Ovo je situacija koja se javlja kod upita koji u ORDER BY klauzuli sadrže kriterijum uređenosti K.

Nad tabelom treba kreirati S-indeks po kriterijumu K.

Implementira: M3

S4 Svođenje rezultata upita po kriterijumu svođenja K

Ovo je situacija koja se javlja kod upita sa GROUP BY klauzulom.

Nad tabelom treba kreirati S-indeks po kriterijumu svođenja K. Ubrzavajući efekat je u tome što se redovima tabele pristupa uređeno po kriterijumu svođenja, tako da se u svakom trenutku formira samo jedan svodni red.

Implementira: M3

S5 Spajanje tabela po mehanizmu strani ključ - primarni ključ

Neka su tabele koje učestvuju u spajanju $R(\underline{IDR}, <OR>)$ i $S(\underline{IDS}, <OS>, IDR)$.

U slučaju spajanja $S > IDR < R$ treba pristupiti svim redovima u S i maksimalno smanjiti broj pristupa redovima u R. To se obezbeđuje pomoću indeksa:
S : S-indeks po IDR, R : H-indeks po IDR (po pravilu već postoji). Ubrzavajući efekat je u tome što se redovima u S pristupa po grupama sa jednakim IDR pa je za celu takvu grupu dovoljno jednom pristupiti orgovarajućem redu u R.

U slučaju spajanja $R > IDR < S$ treba pristupiti svim redovima u R i maksimalno smanjiti broj pristupa redovima u S. To se obezbeđuje pomoću indeksa:
R : H-indeks po IDR (po pravilu već postoji), S : S-indeks po IDR. Ubrzavajući efekat je u tome što se za svako IDR očitano iz R pristupa direktno grupi redova u S sa tom vrednošću IDR.

Implementira: M3 za spajanje $S > IDR < R$,
M2 za spajanje $R > IDR < S$.

Napomene: U slučaju da se referišuća tabela S spaja sa više ciljnih tabela $R_1 \dots R_i$, za S treba uvesti indeks:
S : S-indeks po IDR_1, \dots, IDR_i .

U slučaju spajanja $R > IDR < S$ uvođenje S-indeksa
R : S-indeks po IDR dodatno ubrzava spajanje, pošto je prolaz kroz S u jednom smeru neprekidno.