

# Indeks-sekvencijalna organizacija datoteke

Struktura, formiranje, traženje, obrada, ažuriranje, primena i ocena

# Sadržaj

- ▶ Indeksne datoteke
- ▶ Indeks-sekvencijalna organizacija datoteke
- ▶ Indeks-sekvencijalna metoda pristupa

# Indeksne datoteke

## ► Karakteristike indeksnih datoteka

### ► postojanje **indeksa**

- pomoćna struktura podataka
- realizovana u posebnoj datoteci, kao stablo traženja
- sadrži parove

*(vrednost ključa, relativna adresa sloga/bloka)*

- za preslikavanje argumenta traženja u adresu sloga
- za brz pristup slučajno odabranom i logički narednom slogu

### ► smeštanje kompletnih slogova

- u posebnu datoteku - zonu podataka (primarnu zonu)
- može biti različito organizovana

# Indeksne datoteke

## ► Vrste indeksnih datoteka

### ► **statičke**

- istorijski prve
  - danas ređe u praktičnoj upotrebi
- statička alokacija memorijskog prostora
  - definiše se prilikom projektovanja organizacije
- statički indeks
  - formira se i, nakon formiranja, ne ažurira se

### ► **dinamičke**

- praktično nezaobilazne u realnim projektima
- dinamička alokacija memorijskog prostora
- dinamički indeks
  - ažurira se paralelno sa ažuriranjem zone podataka
  - održavanje poželjnih karakteristika u vremenu

# Indeksne datoteke

- ▶ **Efikasnost organizacija u upotrebi**
  - ▶ sekvencijalne datoteke
    - ▶ ideal redosledne obrade i traženja logički narednih slogova
  - ▶ rasute datoteke
    - ▶ ideal direktne obrade i traženja slučajno odabralih slogova
  - ▶ indeksne datoteke
    - ▶ struktura kompromisa
    - ▶ solidna podrška direktne i redosledne obrade, kao i obe vrste traženja

# Sadržaj

- ▶ Indeksne datoteke
- ▶ Indeks-sekvencijalna organizacija datoteke
- ▶ Indeks-sekvencijalna metoda pristupa

# Indeks-sekvencijalna organizacija

## ► **Statička indeks-sekvencijalna datoteka**

- tri memorijske zone ili osnovne organizacije datoteke
  - primarna zona ili zona podataka
    - sekvensijalna organizacija
  - zona indeksa
    - spregnuta organizacija - n-arno stablo
  - zona prekoračenja
    - spregnuta organizacija - lanci prekoračilaca

# Indeks-sekvencijalna organizacija

## ► Primarna zona

- ▶ slogovi uređeni saglasno rastućim vrednostima ključa
- ▶ slogovi grupisani u blokove
  - ▶ poželjan što veći faktor blokiranja
- ▶ kreira se u postupku formiranja statičke indeks-sekvencijalne datoteke
- ▶ nikada se, naknadno, ne ažurira
- ▶ ciljevi
  - ▶ iskoristiti poželjne osobine sekvencijalne organizacije u redoslednoj obradi podataka
  - ▶ izbeći efekat loših performansi ažuriranja sekvencijalno organizovane datoteke

# Indeks-sekvencijalna organizacija

## ► Zona indeksa

- puno stablo traženja, spregnuta struktura
  - reda  $n$  ( $n \geq 2$ )
  - visine  $h$
- čvor stabla = blok, sadrži od 1 do  $n$  elemenata
  - parovi  $(k_e, A_e)$ ,  $e \in \{1, \dots, n\}$
  - $n$  - faktor blokiranja u zoni indeksa
    - $k_e = k(S)$  - vrednost ključa sloga  $S$
    - $A_e$  je adresa:
      - bloka primarne zone u kojem je slog  $S$ , u slučaju lista, ili
      - podstabla, tj. drugog čvora stabla traženja koji takođe sadrži  $k_e$ , u slučaju neterminalnog čvora
  - elementi u čvoru uređeni saglasno rastućim vrednostima ključa  $k_e$ 
    - čvor je sekvencijalno organizovana struktura

# Indeks-sekvencijalna organizacija

## ► Zona indeksa

### ► retko popunjjeni indeks

- reprezentativne vrednosti ključa svakog bloka primarne zone propagirane u stablo

### ► vrednosti ključa $k_e$ u stablu

- najmanje ili najveće vrednosti ključa iz svakog bloka primarne zone

### ► elementi listova stabla

- sadrže po jednu vrednost ključa iz svakog bloka

### ► elementi čvorova na višim nivoima hijerarhije stabla

- sadrže po jednu vrednost ključa iz svakog direktno podređenog čvora

- vrednosti ključa ponavljaju se u čvorovima na svim nižim nivoima hijerarhije

# Indeks-sekvencijalna organizacija

## ► **Zona indeksa**

- neterminalni čvor sa  $m$  ( $1 \leq m \leq n$ ) elemenata
  - poseduje  $m$  direktno podređenih čvorova
- adresa  $A_e$  - pokazivač, relativna adresa

## ► **Vrste zone indeksa**

- **zona indeksa s propagacijom najvećih vrednosti ključa iz svakog bloka**
  - u slučaju poslednjeg bloka, propagira se ne aktuelna najveća vrednost ključa, već najveća dozvoljena vrednost ključa
- **zona indeksa s propagacijom najmanjih vrednosti ključa iz svakog bloka**
  - u slučaju prvog bloka, propagira se ne aktuelna najmanja vrednost ključa, već najmanja dozvoljena vrednost ključa

# Indeks-sekvencijalna organizacija

- ▶ Primer - indeks-sekvencijalna datoteka  $D_{insek}$ 
  - ▶ propagacija najvećih vrednosti ključa
    - ▶ iz svakog bloka primarne zone, osim iz poslednjeg
    - ▶ najveća dozvoljena vrednost ključa - iz poslednjeg bloka
  - ▶  $N = 13$  slogova
  - ▶ faktor blokiranja u primarnoj zoni  $f = 3$
  - ▶ red stabla traženja i faktor blokiranja u zoni indeksa  $n = 2$

# Indeks-sekvencijalna organizacija

- ▶ Primer - indeks-sekvencijalna datoteka  $D_{insek}$ 
  - ▶ propagacija najvećih vrednosti ključa

PRIMARNA ZONA

$A_1^p$						
03	$p(S_1)$	07	$p(S_2)$	13	$p(S_3)$	

$A_2^p$						
15	$p(S_4)$	19	$p(S_5)$	23	$p(S_6)$	

$A_3^p$						
25	$p(S_7)$	27	$p(S_8)$	29	$p(S_9)$	

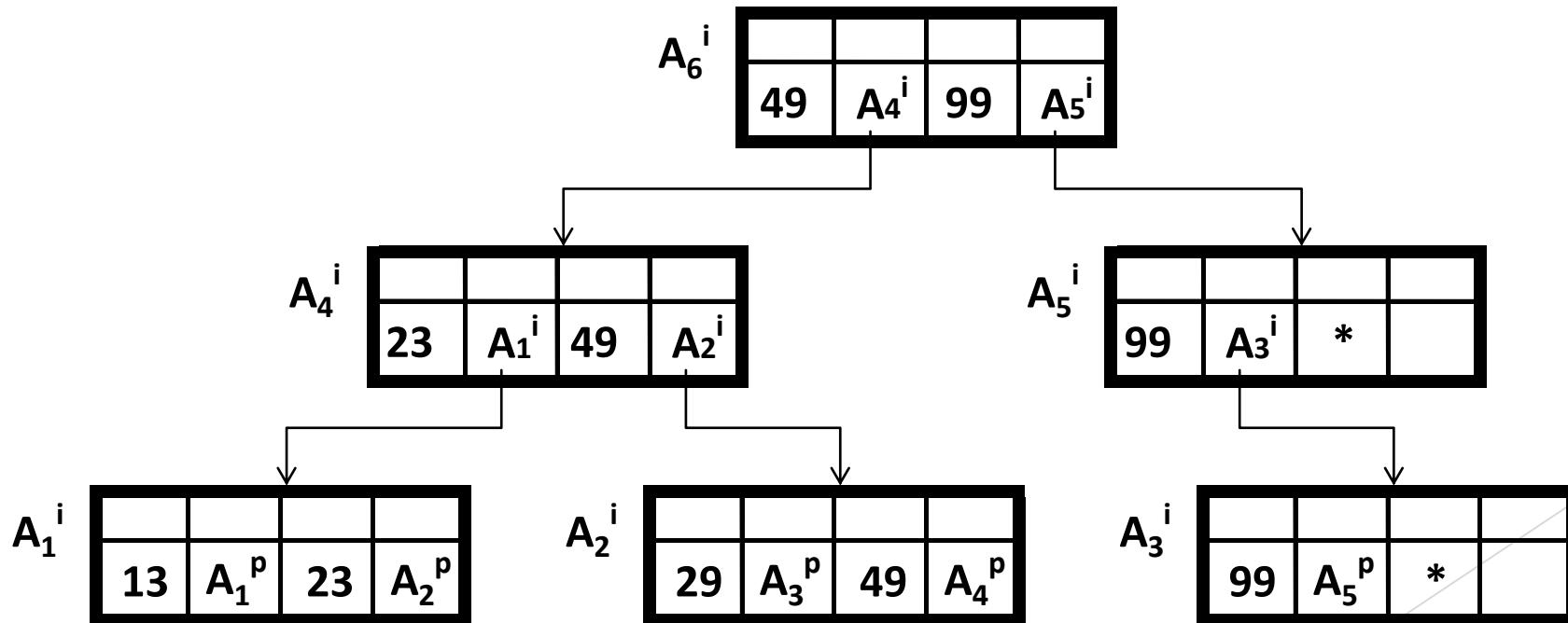
$A_4^p$						
34	$p(S_{10})$	43	$p(S_{11})$	49	$p(S_{12})$	

$A_5^p$						
64	$p(S_{13})$	*				

# Indeks-sekvencijalna organizacija

- ▶ Primer - indeks-sekvencijalna datoteka  $D_{insek}$ 
  - ▶ propagacija najvećih vrednosti ključa

## ZONA INDEKSA



# Indeks-sekvencijalna organizacija

## ► Zona indeksa

- broj čvorova  $C_i$ 
  - na  $i$ -tom nivou hijerarhije stabla ( $i=1, \dots, h$ )

$$C_i = \left\lceil \frac{B}{n^{h-i+1}} \right\rceil$$

►  $B$  - broj blokova u primarnoj zoni ( $B \geq 1$ )

- visina stabla

$$h = \lceil \log_n B \rceil$$

- ukupni broj čvorova stabla  $C$

$$C = \sum_{i=1}^h \left\lceil \frac{B}{n^{h-i+1}} \right\rceil$$

# Indeks-sekvencijalna organizacija

## ► Zona indeksa

- kapacitet stabla
  - koliko parova ( $k_e, A_e$ ) se može upisati u čvorove

$$K = nC$$

- ukupni broj elemenata  $E$  u stablu pristupa

$$E = \sum_{i=1}^h \left\lceil \frac{B}{n^{h-i}} \right\rceil$$

- stablo traženja obezbeđuje relativno brz pristup za traženje slučajno odabranog sloga

# Indeks-sekvencijalna organizacija

## ► Zona prekoračenja

- sadrži kompletne slogove datoteke
  - kao i primarna zona
  - koji se upisuju u zonu prekoračenja pri upisu novih slogova
  - koji se nazivaju prekoračioci
  - svaki blok primarne zone može imati svoje prekoračioce

- posmatra se list u stablu traženja sa  $m$  elemenata  
 $(1 \leq m \leq n)$ :

$$(k_1, A_1), \dots, (k_{e-1}, A_{e-1}), (k_e, A_e), (k_{e+1}, A_{e+1}), \dots, (k_m, A_m)$$

- $A_e, e \in \{1, \dots, m\}$  - adrese sukcesivnih blokova primarne zone
- ako stablo sadrži najveće vrednosti ključa, slog sa  $k(S)$  je u
  - bloku sa adresom  $A_1$  ako je  $k(S) \leq k_1$
  - bloku sa adresom  $A_e$ , ako je  $k_{e-1} < k(S) \leq k_e, e \in \{2, \dots, m\}$

# Indeks-sekvencijalna organizacija

## ► **Zona prekoračenja**

- ▶ kada **blok u primarnoj zoni nije kompletan** ( $m \neq n$ )
  - ▶ upis novog sloga dovodi samo do pomeranja slogova u bloku
- ▶ kada je **blok u primarnoj zoni kompletan** ( $m = n$ )
  - ▶ upis svakog novog sloga dovodi do upisa jednog od slogova koji pripadaju bloku sa adresom  $A_e$ ,  $e \in \{1, \dots, m\}$ , u zonu prekoračenja
  - ▶  $k_{e(max)}$  - trenutno maksimalna vrednost ključa u bloku sa adresom  $A_e$
  - ▶ ako je  $k(S) < k_{e(max)}$ 
    - ▶ novi slog se upisuje u blok a svi slogovi sa vrednošću ključa većom od  $k(S)$  pomeraju se za jednu lokaciju ka kraju bloka
    - ▶ slog sa vrednošću ključa  $k_{e(max)}$  se upisuje u zonu prekoračenja
  - ▶ ako je  $k(S) > k_{e(max)}$ 
    - ▶ novi slog se upisuje u zonu prekoračenja

# Indeks-sekvencijalna organizacija

## ► **Zona prekoračenja**

- sprežu se logički neposredno susedni prekoračioci iz jednog bloka, faktor blokiranja  $f_z = 1$ 
  - za svaki blok primarne zone, najviše jedan lanac spregnutih prekoračilaca
  - slogovi u svakom lancu prekoračilaca uređeni su u rastućem (alternativno opadajućem) redosledu
  - dodatno, lanac slobodnih blokova

## ► Pokazivač na početak lanca - dva načina

- **direktno povezivanje sa listom stabla traženja**
  - pokazivač na početak lanca smešta se u odgovarajući list
- **indirektno povezivanje sa listom stabla traženja**
  - pokazivač na početak lanca smešta se u prateći deo odgovarajućeg bloka u primarnoj zoni

# Indeks-sekvencijalna organizacija

- ▶ Struktura zone prekoračenja - direktni pristup

LISTOVI STABLA TRAŽENJA

$A_1^i$	13	$A_1^p$	13	$A_1^p$	19	$A_2^p$	23	$A_2^z$

$A_2^i$	29	$A_3^p$	29	$A_3^p$	43	$A_4^p$	49	$A_3^z$

$A_3^i$	71	$A_5^p$	99	$A_5^p$	*			

PRIMARNA ZONA

$A_1^p$							
	03	$p(S_1)$	07	$p(S_2)$	13	$p(S_3)$	

$A_2^p$							
	14	$p(S_{15})$	15	$p(S_4)$	19	$p(S_5)$	

$A_3^p$							
	25	$p(S_7)$	27	$p(S_8)$	29	$p(S_9)$	

$A_4^p$							
	31	$p(S_{14})$	34	$p(S_{10})$	43	$p(S_{11})$	

$A_5^p$							
	64	$p(S_{13})$	71	$p(S_{17})$	*		

E	$A_4^z$

$A_1^z$	
	49 $p(S_{12})$

$A_2^z$	
	23 $p(S_6)$

$A_3^z$	
	47 $p(S_{16})$

$A_4^z$	
	$A_5^z$

$A_5^z$	
	*

20

# Indeks-sekvencijalna organizacija

- ▶ Struktura zone prekoračenja - indirektni pristup

PRIMARNA ZONA							
$A_1^p$	01	$p(S_{17})$	03	$p(S_1)$	07	$p(S_2)$	$A_4^z$
$A_2^p$	15	$p(S_4)$	19	$p(S_5)$	23	$p(S_6)$	*
$A_3^p$	25	$p(S_7)$	27	$p(S_8)$	29	$p(S_9)$	$A_1^z$
$A_4^p$	31	$p(S_{10})$	43	$p(S_{11})$	47	$p(S_{16})$	$A_3^z$
$A_5^p$	64	$p(S_{13})$	*				*

## ZONA PREKORAČENJA

E	$A_5^z$
$A_1^z$	
30	$p(S_{14})$
*	
$A_2^z$	
14	$p(S_{15})$
*	
$A_3^z$	
49	$p(S_{12})$
*	
$A_4^z$	
13	$p(S_3)$
$A_2^z$	
	*
$A_5^z$	

# Sadržaj

- ▶ Indeksne datoteke
- ▶ Indeks-sekvencijalna organizacija datoteke
- ▶ Indeks-sekvencijalna metoda pristupa

# Indeks-sekvencijalna metoda pristupa

- ▶ Podržana
  - ▶ najčešće, sistemima za upravljanje datoteka, ugrađenim u OS mainframe računara
  - ▶ ređe savremenim SUBP-ovima
- ▶ Obezbeđuje
  - ▶ formiranje, traženje, ažuriranje i reorganizaciju
  - ▶ sekvencijalni, direktni i dinamički način pristupa indeks-sekvencijalnoj datoteci

# Indeks-sekvencijalna metoda pristupa

## ► Formiranje IS datoteke

- ▶ program redosledno učitava slogove ulazne sekvencijalne datoteke
- ▶ smešta blokove u primarnu zonu IS datoteke
- ▶ alternativno, već formirana sekvencijalna datoteka proglašava se primarnom zonom IS datoteke
- ▶ formiranje zone indeksa



# Indeks-sekvencijalna metoda pristupa

## ► Formiranje IS datoteke

- ▶ formiranje zone indeksa
  - ▶ iterativan postupak, po nivoima stabla traženja
    - ▶ s dna na gore → s leva na desno
  - ▶ prvo se formiraju svi listovi - čvorovi nivoa  $h$ , zatim čvorovi nivoa  $h - 1$ , itd. do čvorova nivoa 1
  - ▶ u svaki čvor na  $i$ -tom nivou hijerarhije ( $i=h-1, h-2, \dots, 1$ )
    - ▶ upisuju se najveće (alternativno najmanje) vrednosti ključa iz  $n$  sukcesivnih čvorova na nivou hijerarhije  $i+1$
  - ▶ propagacija najvećih vrednosti
    - ▶ u poslednji element krajnjeg desnog čvora upisuje se maksimalna dozvoljena vrednost ključa
  - ▶ propagacija najmanjih vrednosti
    - ▶ u prvi element krajnjeg levog čvora upisuje se minimalna dozvoljena vrednost ključa

# Indeks-sekvencijalna metoda pristupa

## ► **Formiranje IS datoteke**

- formiranje zone prekoračenja
  - alocira se prazna zona prekoračenja
  - svi blokovi sprežu se u lanac slobodnih blokova
  - početak lanca upisuje se u zaglavlje zone prekoračenja

# Indeks-sekvencijalna metoda pristupa

## ► Traženje logički narednog sloga

- ▶ vrši se kombinovanom primenom
  - ▶ metode linearног traženja i
  - ▶ metode traženja praćenjem pokazivača
- ▶ počinje u prvom bloku primarne zone
  - ▶ svako naredno traženje se nastavlja od tekućeg sloga datoteke u bloku primarne zone
    - ▶ linearna metoda
  - ▶ po dolasku do poslednjeg sloga bloka traženje se nastavlja u lancu prekoračilaca, ako postoji
    - ▶ metoda praćenja pokazivača
  - ▶ indirektno povezivanje prekoračilaca - nastavak traženja direktno u zoni prekoračenja
  - ▶ direktno povezivanje prekoračilaca - pristup direktno nadređenom listu i nastavak traženja u zoni prekoračenja

# Indeks-sekvencijalna metoda pristupa

- ▶ **Traženje logički narednog sloga**
  - ▶ direktno povezivanje prekoračilaca
    - ▶ pristupa se blokovima primarne zone, prekoračiocima i listovima stabla traženja
    - ▶ broj pristupa  $R$  i pri uspešnom i pri neuspešnom traženju jednog logički narednog sloga

$$0 \leq R \leq B + \left\lceil \frac{B}{n} \right\rceil + Z - (i + j + k)$$

- ▶  $Z$  - ukupni broj slogova u zoni prekoračenja
- ▶  $i$  - redni broj tekućeg bloka datoteke u odnosu na početak primarne zone
- ▶  $j = \lceil i/n \rceil$  - redni broj tekućeg lista stabla traženja
- ▶  $k$  - broj slogova zone prekoračenja kojima se već pristupilo

# Indeks-sekvencijalna metoda pristupa

- ▶ **Traženje logički narednog sloga**
  - ▶ indirektno povezivanje prekoračilaca
    - ▶ broj pristupa  $R$  i pri uspešnom i pri neuspešnom traženju jednog logički narednog sloga

$$0 \leq R \leq B + Z - (i + k)$$

- ▶ traženje logički narednog sloga je efikasnije

# Indeks-sekvencijalna metoda pristupa

## ► **Traženje slučajno odabranog sloga**

- praćenjem pokazivača u stablu pristupa
  - započinje u korenu i stiže do lista u stablu traženja
  - uvažava organizaciju sa propagacijom maksimalnih ili minimalnih vrednosti ključa iz svakog bloka

## ► **direktno povezivanje prekoračilaca**

- dolaskom do odgovarajućeg elementa u listu - odluka o nastavku traženja
  - u bloku primarne zone ili
  - praćenjem lanca prekoračilaca, u zoni prekoračenja

## ► **indirektno povezivanje prekoračilaca**

- prati se pokazivač odgovarajućeg elementa u listu i prelazi se u blok podataka u primarnoj zoni
- po potrebi, nastavlja se traženje praćenjem lanca prekoračilaca

# Indeks-sekvencijalna metoda pristupa

- ▶ **Traženje slučajno odabranog sloga**
  - ▶ direktno povezivanje prekoračilaca
    - ▶ broj pristupa  $R$  i pri uspešnom i pri neuspešnom traženju jednog slučajno odabranog sloga

$$h+1 \leq R \leq h+z$$

- ▶ indirektno povezivanje prekoračilaca
  - ▶ broj pristupa  $R$  i pri uspešnom i pri neuspešnom traženju jednog slučajno odabranog sloga

$$h+1 \leq R \leq h+1+z$$

- ▶  $z$  - dužina lanca prekoračilaca za jedan blok primarne zone
- ▶ nešto efikasnije traženje u datoteci s direktnim povezivanjem prekoračilaca

# Indeks-sekvencijalna metoda pristupa

## ► Obrada IS datoteke

- moguća efikasna obrada i
  - u režimu redosledne obrade i
  - u režimu direktne obradne
- pogodne za korišćenje u ulozi vodeće datoteke u oba režima
- redosledna obrada putem vodeće datoteke od  $N_v$  slogova odvija se
  - naizmeničnim pristupanjem blokovima primarne zone i njihovim lancima prekoračilaca
  - adresa prvog bloka primarne zone nalazi se u zaglavlju datoteke

# Indeks-sekvencijalna metoda pristupa

## ► Obrada IS datoteke

- ukupan broj pristupa pri redoslednoj obradi  $R_{uk}$

- za slučaj direktnog povezivanja

$$R_{uk} = B + Z + \lceil B/n \rceil$$

- za slučaj indirektnog povezivanja

$$R_{uk} = B + Z$$

- očekivani broj pristupa pri uspešnom ili neuspešnom traženju jednog logički narednog sloga

- za slučaj direktnog povezivanja

$$\bar{R} = \frac{B + Z + \lceil B/n \rceil}{N_v}$$

- za slučaj indirektnog povezivanja

$$\bar{R} = \frac{B + Z}{N_v}$$

# Indeks-sekvencijalna metoda pristupa

- ▶ **Obrada IS datoteke**
  - ▶ redosledna obrada
    - ▶ nešto efikasnija kod datoteke sa indirektnim povezivanjem prekoračilaca
    - ▶ pri uobičajenim vrednostima reda stabla  $n$  razlika je neznatna

# Indeks-sekvencijalna metoda pristupa

- ▶ **Obrada IS datoteke**

- ▶ **direktna obrada**

- ▶ putem vodeće datoteke od  $N_v$  slogova

$$N_v = N_v^u + N_v^n$$

- ▶ očekivani ukupni broj pristupa

$$\bar{R}_{uk} = \bar{R}_u N_v^u + \bar{R}_n N_v^n$$

- ▶ - očekivani broj pristupa pri uspešnom traženju

$$\bar{R}_u$$

- ▶ - očekivani broj pristupa pri neuspešnom traženju

$$\bar{R}_n$$

# Indeks-sekvencijalna metoda pristupa

## ► Obrada IS datoteke

### ► direktna obrada

- putem vodeće datoteke od  $N_v$  slogova
- očekivani broj prekoračilaca po bloku primarne zone

$$\bar{z} = \frac{Z}{B}$$

$$\bar{R}_u = \bar{R}_n = h + \frac{f}{f + \bar{z}} + \frac{\bar{z}}{f + \bar{z}} \sum_{i=1}^{\bar{z}} i \frac{1}{\bar{z}}$$

- za datoteku sa direktnim povezivanjem prekoračilaca
  - ista verovatnoća zaustavljanja traženja na bilo kom slogu

$$\bar{R}_u = \bar{R}_n = h + \frac{2f + \bar{z}(\bar{z}+1)}{2(f + \bar{z})}$$

# Indeks-sekvencijalna metoda pristupa

- ▶ **Obrada IS datoteke**

- ▶ **direktna obrada**

- ▶ putem vodeće datoteke od  $N_v$  slogova
    - ▶ za datoteku sa indirektnim povezivanjem prekoračilaca

$$\bar{R}_u = h + 1 + \frac{\bar{z}}{f + \bar{z}} \sum_{i=1}^{\bar{z}} i \frac{1}{\bar{z}} = h + 1 + \frac{\bar{z}(\bar{z}+1)}{2(f + \bar{z})}$$

$$\bar{R}_n = h + 1 + \frac{\bar{z}(\bar{z}+3)}{2(f + \bar{z})}$$

- ▶ **direktna obrada**

- ▶ nešto efikasnija kod datoteke sa direktnim povezivanjem prekoračilaca

# Indeks-sekvencijalna metoda pristupa

## ► Ažuriranje IS datoteke

- ▶ vrši se u režimu direktne obrade
- ▶ **upis novog sloga**
  - ▶ nakon neuspešnog traženja
  - ▶ ako se neuspešno traženje zaustavilo u bloku primarne zone
    - ▶ vrši se pomeranje slogova sa većom vrednošću ključa od vrednosti ključa novog sloga za jednu lokaciju ka kraju bloka
    - ▶ novi slog se upisuje u lokaciju koju je zauzimao slog sa prvom većom vrednošću ključa, a slog sa do tada najvećom vrednošću ključa u bloku upisuje se u zonu prekoračenja
    - ▶ prekoračilac se upisuje u lokaciju čiju adresu sadrži indeks slobodnih lokacija i povezuje se sa ostalim prekoračiocima iz bloka

# Indeks-sekvencijalna metoda pristupa

- ▶ **Ažuriranje IS datoteke**
  - ▶ upis novog sloga
    - ▶ ako se neuspjeh traženje zaustavilo na nekom od prekoračilaca
      - ▶ novi slog se upisuje u prvu slobodnu lokaciju
      - ▶ uvezuje se sa ostalim prekoračiocima

# Indeks-sekvencijalna metoda pristupa

- ▶ **Ažuriranje IS datoteka**
  - ▶ brisanje sloga
    - ▶ logičko brisanje - češće

$$R_d = R_u + 1$$

- ▶ lokacija logički izbrisanih sloga može se upotrebiti za upis novog sloga u specijalnom slučaju
  - ▶ kada se vrednost ključa novog sloga nalazi tačno u odgovarajućim granicama
- ▶ **fizičko brisanje**
  - ▶ zahteva pomeranje slogova sa ažuriranjem lanca prekoračilaca
  - ▶ zahteva veći broj pristupa datoteci

# Indeks-sekvencijalna metoda pristupa

## ► Ažuriranje IS datoteke

- modifikacija sadržaja postojećeg sloga
  - nakon uspešnog traženja
  - potreban samo jedan pristup da bi se modifikovani slog upisao u datoteku

# Indeks-sekvencijalna metoda pristupa

## ► Reorganizacija IS datoteke

- značajna degradacija performansi traženja slogova i obrade datoteke u vremenu
  - usled upisa slogova u zonu prekoračenja i
  - logičkog brisanja slogova
- periodična reorganizacija datoteke
  - uklanjanje negativnih posledica ažuriranja
- postupak
  - ponovno formiranje primarne zone
    - redoslednom obradom - traženjima logički narednih slogova u postojećoj primarnoj zoni i zoni prekoračenja
  - generisanje novog stabla traženja
  - formiranje nove, prazne zone prekoračenja

# Indeks-sekvencijalna metoda pristupa

## ► Reorganizacija IS datoteke

- interval vremena između dve reorganizacije može biti
  - fiksan
    - npr. jednom mesečno
  - dinamički određen
    - na osnovu stepena popunjenoosti zone prekoračenja
    - npr. kada se zona prekoračenja popuni do 80% svog obima, a dimenzionisana je da primi npr. 10% slogova prim. zone
- distribuirani slobodni prostor
  - ublažava problem degradacije performansi obrade zbog upisa novih slogova
    - blokovi podataka se pri formiranju datoteke popunjavaju samo delimično (npr. 60% ili 80%)
    - time se obezbeđuje prostor za upis novih slogova
  - produžava se interval vremena između dve reorganizacije

# Indeks-sekvencijalna metoda pristupa

## ► **Reorganizacija IS datoteke**

- IS datoteka može se ažurirati i kao sekvencijalna u režimu redosledne obrade
  - po završetku ažuriranja, generiše se novo stablo traženja
  - ovakav postupak predstavlja istovremeno i reorganizaciju datoteke

# Indeks-sekvencijalna metoda pristupa

## ► Oblasti primene i ocena karakteristika

- prednosti
  - kada iste podatke treba obrađivati i u režimu redosledne i u režimu direktne obrade
  - intenzivno se koristi u paketnoj obradi
  - može se koristiti i u interaktivnoj obradi
- performanse redosledne obrade
  - u početku ne zaostaju za performansama redosledne obrade sekvencijalne datoteke
- performanse direktne obrade
  - ne zaostaju značajnije za performansama direktne obrade rasute datoteke

# Indeks-sekvencijalna metoda pristupa

## ► Oblasti primene i ocena karakteristika

- glavni nedostatak
  - upis slogova u zonu prekoračenja dovodi do degradacije performansi obrade
  - performanse redosledne obrade mogu se držati pod kontrolom
    - pogodnim dimenzionisanjem zone prekoračenja
  - performanse traženja slučajno odabranog sloga i direktne obrade značajnije se degradiraju
  - jedino rešenje: periodično reorganizovanje datoteke
    - nepogodno, ako se mora često sprovoditi u slučaju datoteka velikog obima

# Indeks-sekvencijalna metoda pristupa

## ► Oblasti primene i ocena karakteristika

- ▶ FSP zasnovane na statičkoj IS organizaciji
  - ▶ intenzivno korišćene u mrežnim sistemima baza podataka
  - ▶ relacioni i objektno-relacioni SUBP ih retko podržavaju
- ▶ ponekad se ističe više njihov istorijski značaj
  - ▶ nazivaju se i klasičnim indeks-sekvencijalnim datotekama
  - ▶ preteča modernih indeksnih datoteka s B-stablima
- ▶ osnovna ideja za primenu IS datoteka
  - ▶ kada se podaci ne ažuriraju intenzivno i u većem obimu
  - ▶ kada je potrebno obezbediti vrlo efikasnu redoslednu obradu i, u isto vreme, solidne performanse direktnе obrade
  - ▶ brz pristup slučajno odabranom slogu u sekvencijalnoj strukturi vrši se korišćenjem stabla traženja kao funkcije koja preslikava vrednost ključa u adresu

# Sadržaj

- ▶ Indeksne datoteke
- ▶ Indeks-sekvencijalna organizacija datoteke
- ▶ Indeks-sekvencijalna metoda pristupa

# Pitanja i komentari



Kraj prezentacije

# Indeks-sekvencijalna organizacija datoteke

Struktura, formiranje, traženje, obrada, ažuriranje, primena i ocena