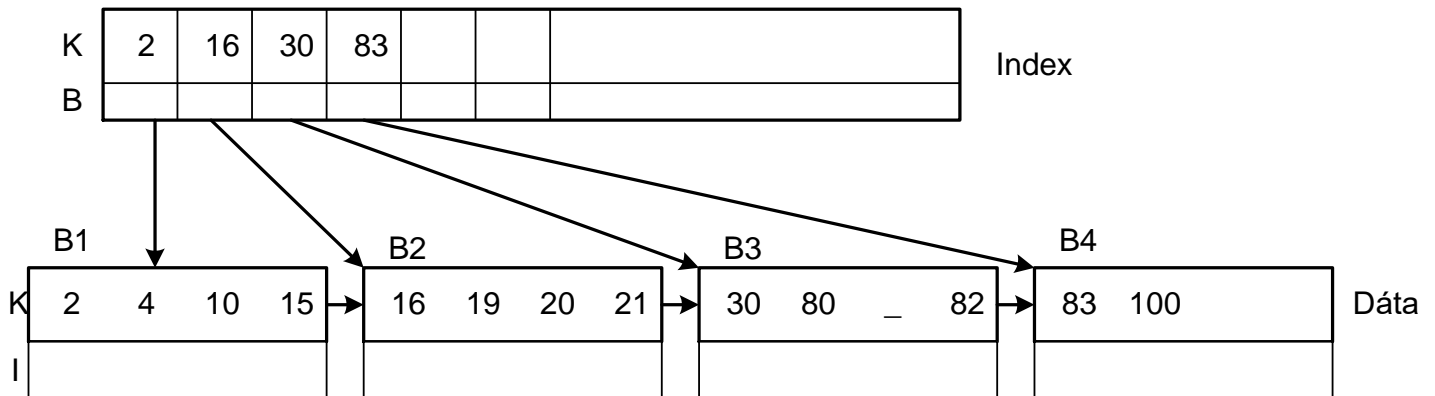


# Indexovaný utriedený súbor

(indexsekvenčný)

## Organizácia:



**Pozostáva z dvoch súborov: Index, Dáta**

**Hlavný súbor (Dáta) utriedený (bloky môžu/nemusia byť zret'azené)**  
**Index: utriedený súbor.**

každý jeho záznam charakterizuje 1 blok hlavného súboru pomocou adresy bloku a najmenšieho kľúča v ňom.

Ak v indexe nasledujú za sebou kľúče  $K_1, K_2$  tak hovoríme, že kľúč  $K_1$  v indexe "pokrýva" všetky kľúče z intervalu  $< K_1, K_2$ ). (30 je pokrytím kľúčov z  $< 30, 83$ ).

**Index všeobecne: blokovaný súbor s prístupom podľa poradia**  
snaha : čo najväčšie bloky (záznamy sú krátke), príp. 1 blok  
**Index sa implementuje ako súvislý utriedený súbor.**

**Pri extrémne veľkých súboroch nie je problém implementovať index indexu!**

**V prípade, ak je súbor Dáta usporiadaný iba cez adresy blokov (nesúvislý utriedený súbor), je vkladanie záznamov podstatne rýchlejšie, keďže nie je potrebné posúvať všetky záznamy s väčšími kľúčmi. Alokujeme len nový blok a ten sa zret'azí, to znamená prenos iba dvoch blokov dát.**

## Spracovanie:

- sekvenčné v poradí kľúčov (rôzne techniky pre zret'azené/nezret'azené bloky)
- priame (vyhľadanie podľa kľúča)

## Priame spracovanie:

- Operácia nájdí záznam s kľúčom K:
  - Nájdí v Indexe B blok, ktorý K pokrýva (modifikáciou binárneho alebo interpolačného vyhľadávania)
  - Prehľadaj blok
  - Ak nenájdenny, prehľadaj preplňujúce bloky (iba pri stratégii 2)

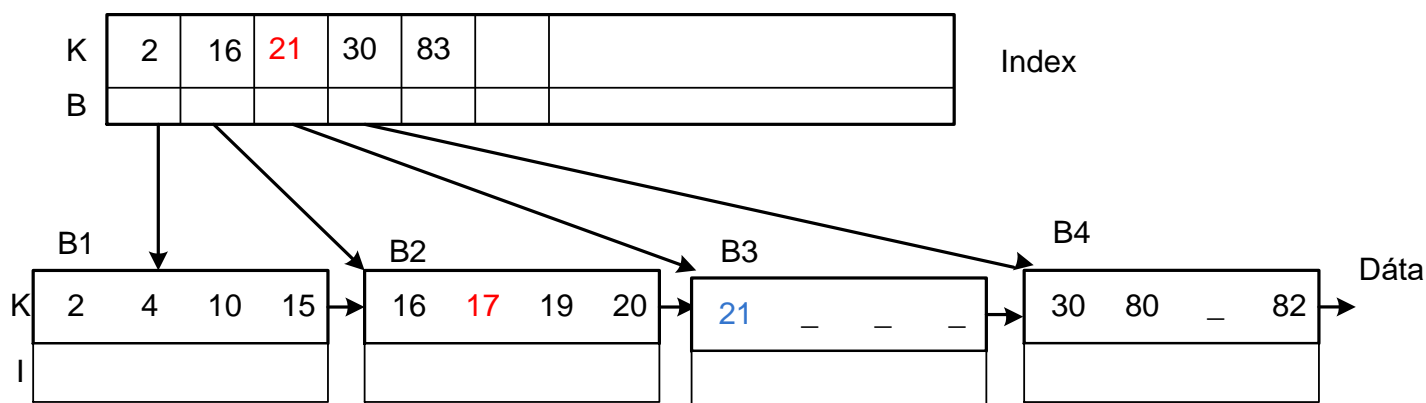
**Nevýhoda:** Častý prenos z bloku do bloku pri dynamických operáciách, preto počet prenosov v najlepšom prípade  $\log_2(\log_2 n) + 1$  (záleží na veľkosti indexu).  
Nepodporuje fixovanie záznamov.

**Výhoda:** Možnosť sekvenčného spracovania, univerzálnosť.

Pri veľkom indexe je lepší samoorganizujúci sa variant: B+ strom.  
Možnosť automatického spustenia reorganizácie. Treba vytvoriť kritérium pre spustenie reorganizácie (hustota, počet preplňujúcich blokov a pod.)

## Vkladanie záznamov:

### Stratégia 1: (dynamická veľkosť indexu)

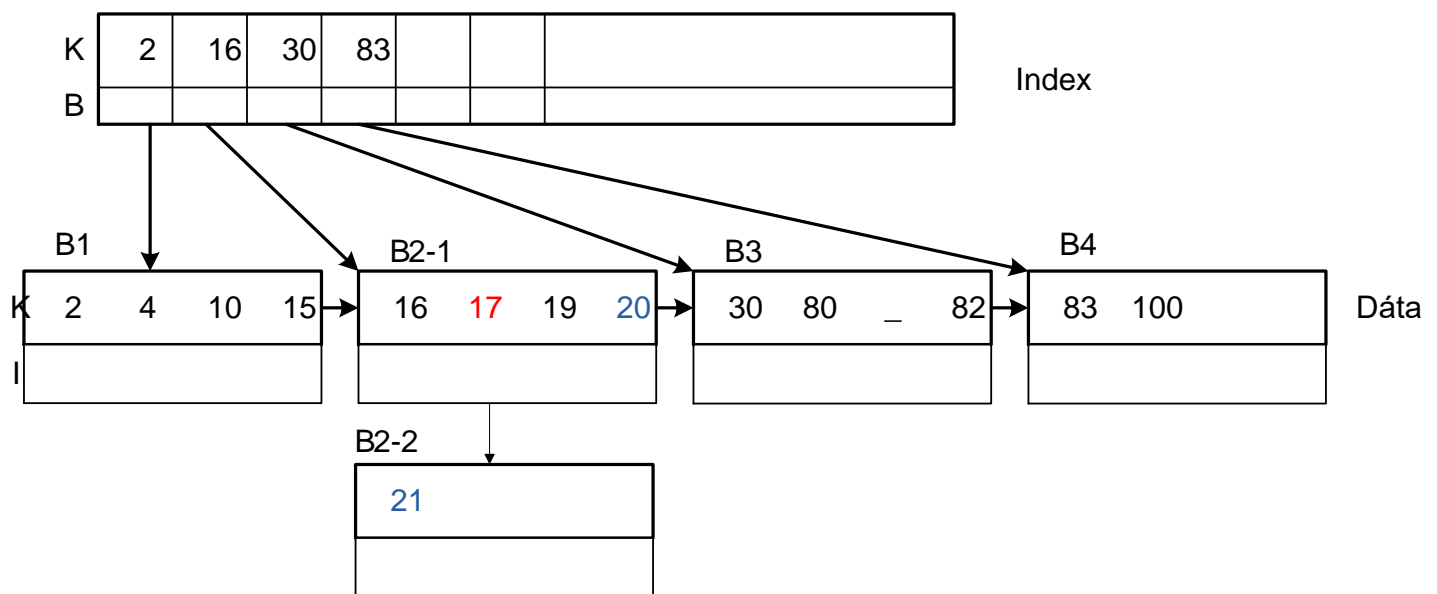


Ak je blok do ktoré sa má vložiť záznam plný, tak sa alokuje nový blok, do ktorého sa presunie najväčší záznam z plného bloku, prípadne vkladany záznam (ak sú kľúče všetkých záznamov v bloku menšie). V príklade na obrázku vloženie čísla 17 vytlačilo 21 do novovzniknutého bloku.

Po úspešnom vložení do súboru Dáta sa upraví príslušné hodnoty v indexe.

Pri mazaní nie je väčšinou vhodné uvoľnené miesto zapíňať. V prípade, že sa vyprázdnil celý blok, ten sa uvoľní a upraví sa aj príslušný index.

## Stratégia 2: (indexovanie skupiny blokov)



Ak je príslušný blok plný, tak sa alokuje preplňujúci blok.