

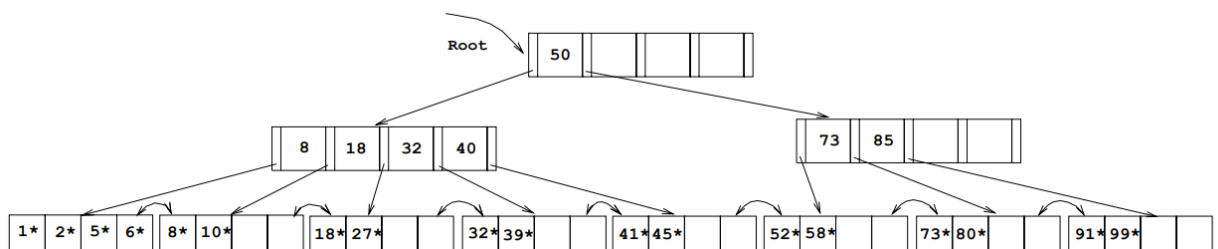
Termin 4. Uređeni indeksi

- Obeleži tačna tvrđenja:
 - Sekundarni indeks može biti klasterišući.
 - Primarni indeks ne može biti redak.
 - Sekundarni indeks mora biti gust.**
 - B+ stablo je višenivojski indeks.**
 - Uređeni indeks podrazumeva da su svi ključevi u indeksu uređeni i svi slogovi u fajlu koji je indeksiran uređeni prema indeksiranim atributu/ima.
 - Glavni negativni efekat koji svaki dodatni indeks nad tabelom ima na performanse baze podataka prilikom operacija dodavanja i brisanja podataka je:
 - Poboljšanje efikasnosti upita, ali ne i umetanja.
 - Sporija pretraga zbog previše indeksa.
 - Dodatno CPU vreme i I/O operacije u slučaju dodavanja ili brisanja slogova.**
 - Povećani prostorni režijski troškovi.
 - Neka tabela ima indeks nad primarnim ključem i indekse nad pojedinim neključnim atributima. Ažuriranja nad tabelom:
 - uvek uključuju ažuriranje svih indeksa
 - ne moraju uključiti ažuriranja bilo kog indeksa**
 - češće uključuju ažuriranje sekundarnih, nego primarnog indeksa nad primarnim ključem**
 - Kolika bi bila zauzetost svakog lista B+-stabla ako bi se unosi indeksa ubacivali sortirano?

Rešenje:

Ako se ključevi dodaju u rastućem redosledu, nove vrednosti se usmeravaju ka poslednjem listu. Kada se ovaj list popuni, on se deli na dva. Od dva čvora nastala deljenjem, levi čvor ostaje netaknut, dok se umetanja nastavljaju u desnom čvoru. Ovo rezultira zauzetošću listova od oko 50 procenata, sa izuzetkom poslednjeg lista.

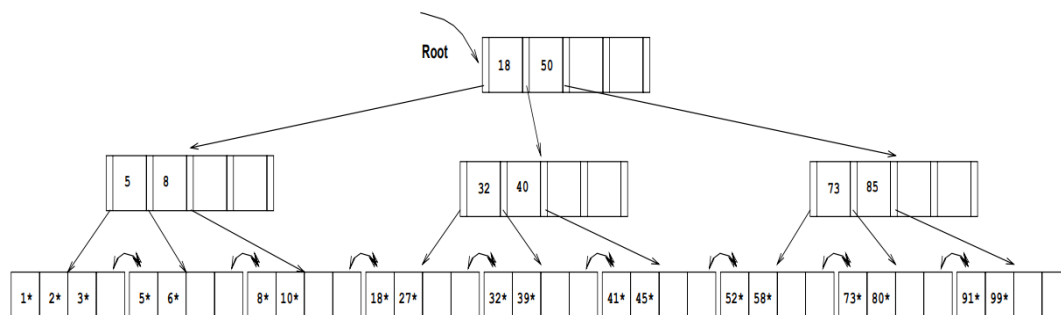
Ako se ključevi dodaju sortirani u opadajućem redosledu, zauzetost bi bila oko 50 procenata za sve čvorove osim prvog lista.
5. Neka je dato sledeće B+ stablo



- Kako će izgledati stablo kada se doda ključ 3?
- Koliko I/O operacija će biti izvršeno?

Rešenje:

a)



b) 4 čitanja (3 do lista i 1 provera sledbenika lista koji mora da distribuira deo ključeva rođaku ili da se podeli) i 6 pisanja (2 nove strane, kore, unutrašnji i 2 lista, podeljeni i rođak)

6. Nacrtaj B+ stablo stepena 4, ako se ključevi dodaju sledećim redom

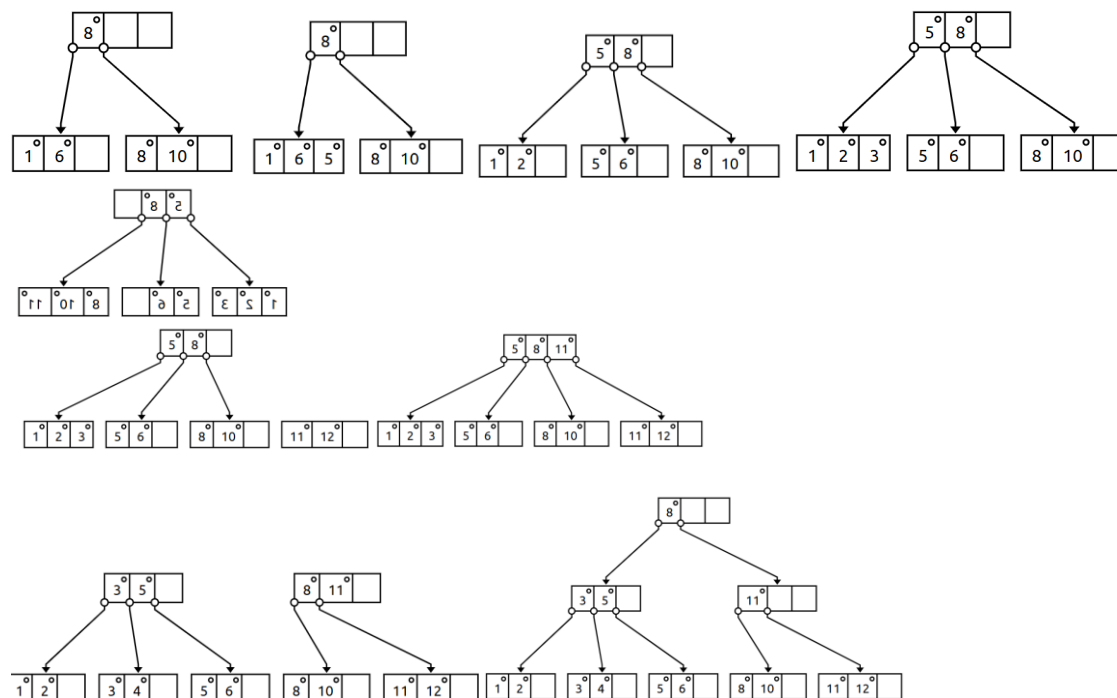
1,6,10,8,5,2,3,11,12,4,14,7,9,13

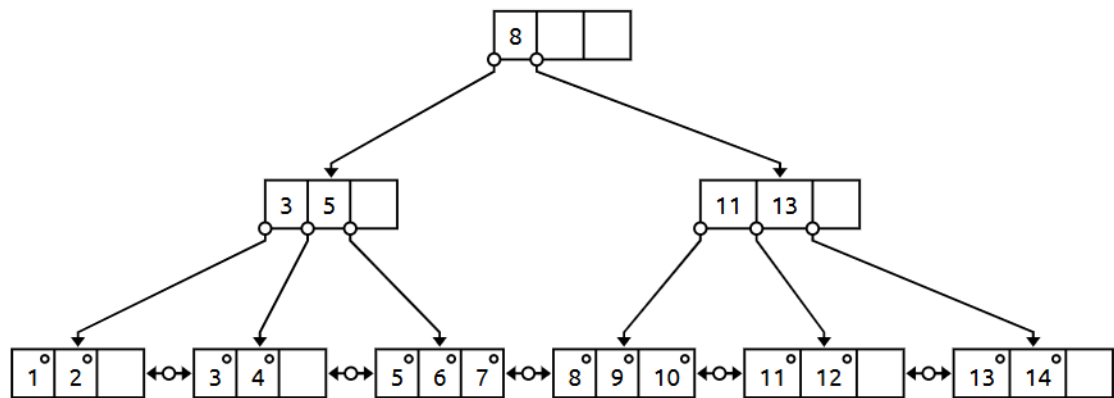
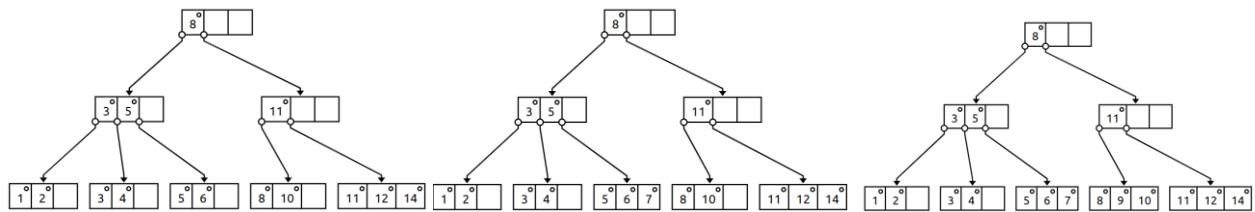
- Kojem se ključevi dodaju jedan po jedan, a čiji se listovi popunjavaju do 100%
- Koje se formira prema bulk load algoritmu opisanom na predavanjima, sa listovima koji se popunjavaju do 75%
- Kako će izgledati stablo dobijeno pod a) brisanjem ključa 5
- Kako će izgledati stablo dobijeno pod b) brisanjem ključa 5

Za crtanje možete koristiti: <https://projects.calebevans.me/b-sketcher/>

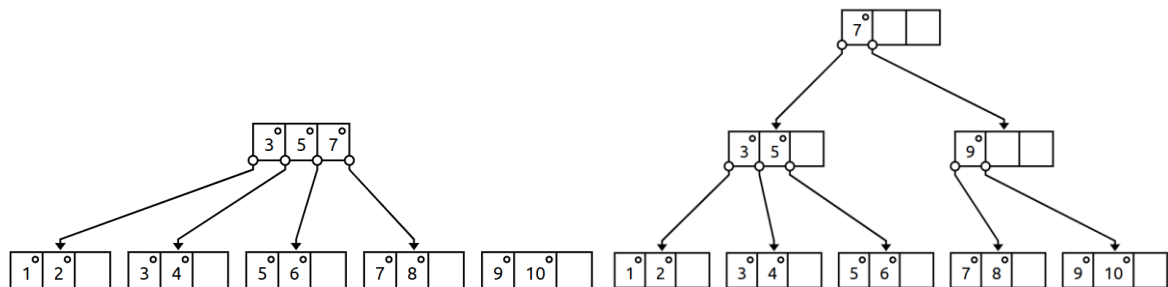
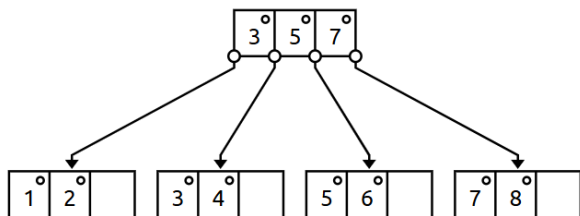
Rešenje:

a)

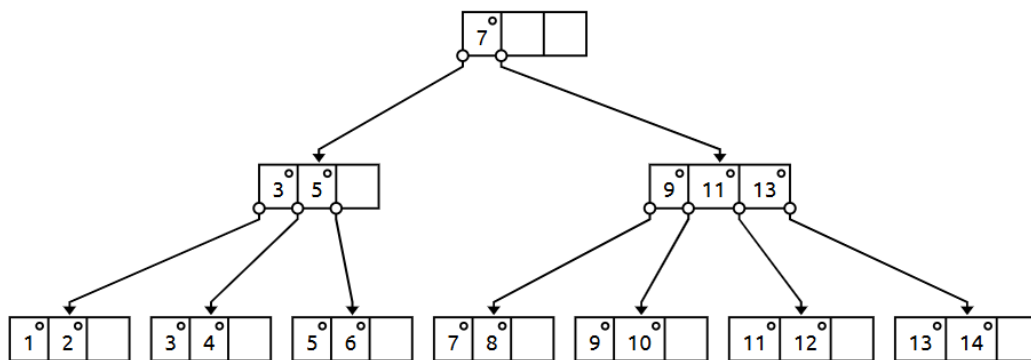




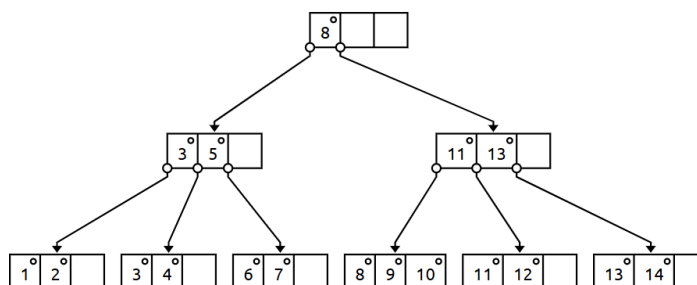
b)



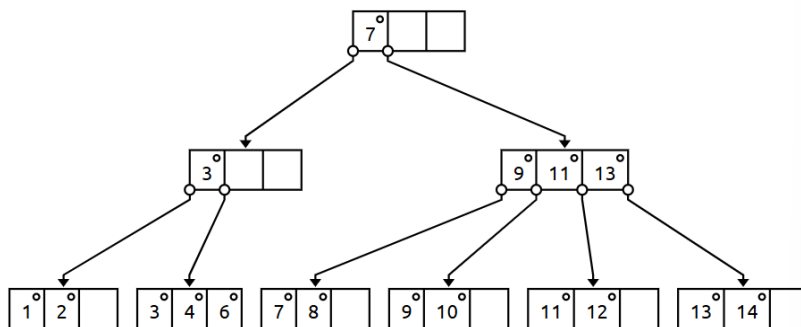
Konačno:



c) Napomena: Ne uklanja se vrednost 5 iz roditelja



d) cxyv



7. Neka je nad heap fajlom od 20000 zapisa napravljen neklasterišući indeks implementiran kao B+ stablo. Neka je veličina ključa za ovo B+ stablo dužine 40 bajtova i predstavlja kandidat ključ (tj. jedinstven je), a pokazivači (tj. ID evi zapisa ili stranica) su (najviše) veličine 10 bajtova. Veličina jedne stranice na disku je 1000 bajtova. Pretpostavimo da su svi čvorovi na svakom nivou su popunjeni koliko god je moguće.
- Koja je visina stabla
 - Koliko čvorova ima na svakom nivou
 - Koji bi kapacitet stabla bio ako ima visinu 4, i popunjenost svih strana ne veću od 70%?

Rešenje:

a) 3

Objašnjenje: neklasterišući, bez ponavljanja vrednosti ključa -> u listovima su isto parovi ključ/vrednost, ne mogu biti celi slogovi

$$N \cdot (40 + 10) + 10 = 1000$$

$$50N = 990 \rightarrow 19 \text{ parova KV, 50 pokazivača}$$

Jedna strana može da sadrži 19 KV parova. Stepen stabla je 20.

Visina stabla:

$$\lceil \log_{20} 20000 \rceil = \left\lceil \frac{\log 20000}{\log 20} \right\rceil = \lceil 3.3 \rceil = 4$$

U visinu ne ulazi koren, pa je visina 3.

b) 1053, 53, 3, 1

Broj listova: $\text{ceil}(20000 \text{ slogova} / 19 \text{ ključeva}) = 1052,63 \rightarrow 1053 \text{ strana}$

Broj čvorova na 2. nivou: $\text{ceil}(1053 / 20) = 53$

Broj čvorova na 1. nivou: $\text{ceil}(53 / 20) = 3$

Koren - 1

c) Sa 70% popunjenosti strana može da prihvati: $\text{floor}(19 * 0.7) = 13$

Maksimalni kapacitet: 13^4