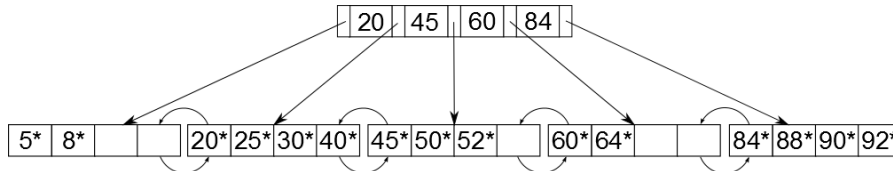


1. Napíšte, pre vami vybrané 2 typy RAIDu, aké bude zrýchlenie zápisu a čítania oproti zápisu na jeden disk, ak použijeme 6 rovnakých diskov.

2. Porovnajcie operáciu vymazania nad súbormi typu Heap, Sorted file, B+strom a Hashovaný súbor z pohľadu výhodnosti. Zdôvodnite svoje tvrdenie.

3. Napíšte kedy je výhodnejšie použiť statické a kedy dynamické indexovanie a prečo.

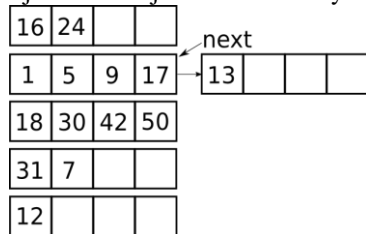
4. Majme B+strom:



Vložte do tohto stromu záznam 65* a následne záznam 85*. Nakreslite situáciu po vložení (aspoň zmenené uzly).

5. Popíšte aspoň dva spôsoby uloženia dát v B+ strome s duplicitným kľúčom (nie UNIQUE stĺpec).

6. Majme nasledujúci index metódy Lineárneho hashovania



Nakreslite index po vložení záznamu s hash hodnotou 6.

7. Máte tabuľku obsahujúcu 100 000 záznamov na 1000 stránkach a buffer, do ktorého vojde 50 stránok. Odhadnite počet I/O operácií na zotriedenie takejto tabuľky za predpokladu, že tabuľka nie je príslušne zotriedená.

8. Predpokladajme, že máme reláciu R, ktorá zaberá M stránok a reláciu S, ktorá zaberá N stránok. Popíšte okolnosti, za ktorých je výhodnejšie využiť na join relácií R a S index join (lebo štandardné block nested loops, hash join a sort merge join sú pomalšie).

9. Majme reláciu človek(id: integer, meno:string, vek:integer, počet_detí:integer), klastrovaný B+ strom na id, a neklastrovaný hash index na meno. Príslušná tabuľka obsahuje 50000 záznamov a zaberá 1000 stránok. Počet kľúčov pre meno 200 a pre počet detí 15. O počte detí vieme z histogramu nasledovné rozloženie:

hodnota	0	1	2	3 – 4	5 – 9	10 – 14
počet	25000	15000	5000	4000	800	200

a, Akú veľkosť výsledku by odhadla databáza pre podmienky vek <5, počet_detí > 10.

b, Ako by ste efektívne počítali nasledujúci dopyt. Napíšte postupnosť operácií, zdôvodnite.

```
SELECT vek, AVERAGE(počet_detí) FROM človek
WHERE (meno="Jano" or meno="Jana")
GROUP BY vek
```

10. Napíšte, ako sa pri nastavení isolation level na read committed a kontrole konkurencie cez viaceré verzie s časovými pečiatkami zabezpečí, že jedna transakcia neprečíta necommitnuté dáta iných transakcií. Ktorá z verzií sa číta?

11. Napíšte, čo sa deje počas fázy analýzy v ARIES algoritme zotavenia (recovery) po výpadku DBMS

12. Porovnajcie základné spôsoby delenia dát pre paralelné a pre distribuované databázy. Prečo sa používajú pri paralelných DB iné stratégie ako pri distribuovaných DB?

13. Popíšte metódu bloom-join.