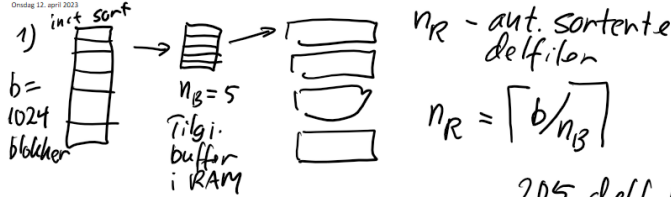


Flettesortering - eksempel

Onsdag 12. april 2023



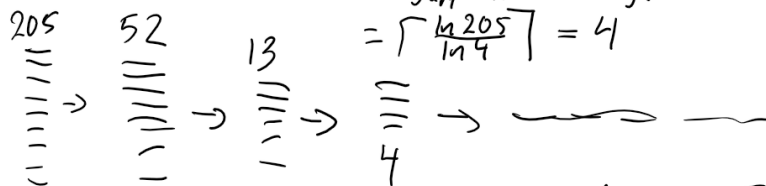
2) Fletting

205 delfiler

Kan flette 4 delfiler om gangen. (flettegrad)

$d_M = \min(n_R, n_B - 1) = 4$

flettepass = $FP = \lceil \log_{d_M}(n_R) \rceil = \lceil \log_4(205) \rceil$



Total I/O : Fase 1 + Fase 2

$1024 \cdot 2$ (les skriv)

$+ 1024 \cdot 4 \cdot 2$ (FP les skriv)

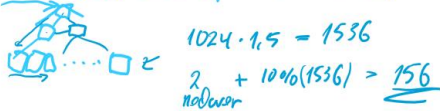
10240 blokker.

SELECT *
FROM employee
WHERE salary < 1000;

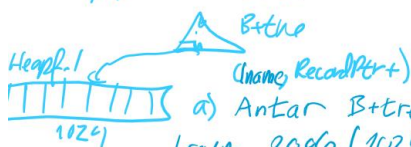
- Antar 10% av postene tilfredsstill betingelsen
- 1024 blokker i en heap
- 200 poster i hver blokk
- Totalt 204.800 poster

1. Scan heapfil 1024 blokker

2. Clustered B+tre på søkeattributt



3. Heapfil + unclustered B+tre



a) Antar B+tre er 20% av $1024 \cdot 1,5$

Lev. 20% $(1024 \cdot 1,5) = 308$ blokker

B+tre 2 nodene + 31 blokker = 33 blokker

b) Følg RecordPtr. =

100% av ant employees $\Rightarrow 20480$ blokker

Potensielt : $33 + 20480 = 20513$ blokker.

Oppgave 5 - Lagring og queryutføring (15 %)

Anta følgende tabell:

CREATE TABLE Student (studno INT, lastname CHAR(30), firstname CHAR(30), email CHAR(30), startyear INT);

Anta tabellen er lagret i en heapfil med 1000 blokker hvor hver blokk får plass til 20 studentposter. Videre er det laget en unclustered B+-tre-indeks på 'lastname'. Vi antar B+-treet har 500 blokker på løv nivå og har høyde 3.

- a) (10 %) Gjør et estimat på hvor mange blokker som aksesseres (leses) ved de følgende SQL-setningene:
- INSERT INTO Student VALUES (12123, 'Hansen', 'Hans', 'hans@email.org', 2013);
 - SELECT lastname, firstname, email, startyear FROM Student WHERE lastname = 'Hansen';
 - SELECT * FROM Student;
 - SELECT DISTINCT lastname FROM Student ORDER BY lastname;

Gi en begrunnelse/forklaring for hvert svar.

- b) (5 %) Hvis den samme tabellen hadde fått kravet om at 'studno' skal være PRIMARY KEY, hvordan ville du ha lagret tabellen? Begrunn svaret.



(lastname, Record Ptr)

a) i) Heap : 1 read + 1 write
B+tre : 3 read + 1 write
= 6 reads

ii) B+tre : 3 read

Heapfil 1 read per 'Hansen'

iii) Les heapfil 1000

iv) Index-only query: $2 + 500 = 502$

b) Index på studno: clustered eller unclustered.

Join - eksempel

Onsdag 12. april 2023

Department 500 poster
lagret i
10 blokker

Employee 60000 poster
2000 blokker

Minne/RAM 7 blokker

1: nested loop

Bruker 5 buffer blokker til den ene
1 buffer til den andre
tabeller.

Total I/O: 1 buffer til resultat

$$2(5 + 2000) = 4010$$

$$\text{Den andre rekkefølgen } 400(5 + 10) = 6000$$

I oppgaven ble det gitt feil svar i utregningen, men riktig i figuren under: 150300.

Oppgave 7 - Join (5 %)

Vi har to tabeller lagret som heapfiler, *Passeringer* og *Kjøretøy*, som skal joines i et query. *Passeringer* består av 15 000 blokker og *Kjøretøy* av 300 blokker. Regn ut I/O-volumet i antall blokker lest ved en nested-loop-join av de to tabellene når det er plass til 32 blokker samtidig i buffer.

>

32 plasser i buffer

30 til Kjøretøy

1 til Passeringer

1 til Resultat

$$10 \cdot (30 + 15000) = \underline{\underline{150300}}$$

Feil
svar var (største tabell)
 $500(30+300) = \underline{165000}$
i RAM)