

Esercizio: Costo ricerca in termini di numero accessi

- Si supponga un file ordinato su campo ID con $r=300.000$ record memorizzati su disco con blocco $B=4096$ di lunghezza. I record hanno lunghezza fissa $R=100$ byte e sono memorizzati in modo unspanned.
- Si assuma un indice primario su campo ID dove:
 - Dimensione ID è 9 byte
 - Dimensione puntatore è 6 byte

Mediamente, con un indice multilivello su ID quanti accessi sono richiesti per una ricerca sul campo ID?

Esercizio: Costo ricerca in termini di numero accessi

- Si supponga un file ordinato su campo ID con $r=300.000$ record memorizzati su disco con blocco $B=4096$ di lunghezza. I record hanno lunghezza fissa $R=100$ byte e sono memorizzati in modo unspanned.
- Si assuma un indice primario su campo ID dove:
 - Dimensione ID è 9 byte
 - Dimensione puntatore è 6 byte

Mediamente, con un indice multilivello su ID quanti accessi sono richiesti per una ricerca sul campo ID?

- Fattore di blocco per il file dati $\lfloor 4096/100 \rfloor = 40$
- Numero blocchi per memorizzare i record: $\lceil 300.000/40 \rceil = 7500$
- Fattore di blocco per il file indice: $\lfloor 4096/15 \rfloor = 273$ **FAN OUT**
- Indice primo livello per 7500 blocchi di dati
 - Numero blocchi per memorizzare indice primo livello: $\lceil 7500/273 \rceil = 28$
- Indice secondo livello per 28 blocchi dell'indice primo livello
 - Numero blocchi per memorizzare indice secondo livello $\lceil 28/273 \rceil = 1 \leq$ blocco radice
- N. Accessi per ricercare record $2+1$

Esercizio: Costo ricerca in termini di numero accessi

- Si supponga un file **NON ordinato** su campo ID con $r=300.000$ record memorizzati su disco con blocco $B=4096$ di lunghezza. I record hanno lunghezza fissa $R=100$ byte e sono memorizzati in modo unspanned.
- Si assuma un indice secondario su campo ID dove:
 - Dimensione ID è 9 byte
 - Dimensione puntatore è 6 byte

Mediamente, con un indice multilivello su ID quanti accessi sono richiesti per una ricerca sul campo ID?

Esercizio: Costo ricerca in termini di numero accessi

- Si supponga un file **NON ordinato** su campo ID con $r=300.000$ record memorizzati su disco con blocco $B=4096$ di lunghezza. I record hanno lunghezza fissa $R=100$ byte e sono memorizzati in modo unspanned.
- Si assuma un indice secondario su campo ID dove:
 - Dimensione ID è 9 byte
 - Dimensione puntatore è 6 byte

Mediamente, con un indice multilivello su ID quanti accessi sono richiesti per una ricerca sul campo ID?

- Fattore di blocco per il file dati $\lfloor 4096/100 \rfloor = 40$
- Numero blocchi per memorizzare i record: $\lceil 300.000/40 \rceil = 7500$
- Fattore di blocco per il file indice: $\lfloor 4096/15 \rfloor = 273$ **FAN OUT**
- Indice primo livello per 300.000 record
 - Numero blocchi per memorizzare indice primo livello: $\lceil 300.000/273 \rceil = 1099$
- Indice secondo livello per 1099 blocchi del primo livello
 - Numero blocchi per memorizzare indice secondo livello $\lceil 1099/273 \rceil = 5$
- Indice terzo livello per 4 blocchi del secondo livello
 - Numero blocchi per memorizzare indice terzo livello $\lceil 5/273 \rceil = 1 \leq$ blocco radice
- N. Accessi per ricercare record $3+1$

Esercizio-1

Si consideri un disco con dimensione di blocco $B = 512$ byte. Un puntatore a blocco è lungo $P = 6$ byte. Il file ha $r = 30.000$ record della tabella IMPIEGATO di *lunghezza fissa*. Ciascun record ha i seguenti campi:

- NOME (30 byte),
- CF (16 byte),
- NUMERO_DIPARTIMENTO (9 byte),
- TELEFONO (9 byte),
- DATA_NASCITA (8 byte),
- CODICE_LAVORO (4 byte).
- Un ulteriore byte viene utilizzato come indicatore di cancellazione.

- Calcolare la dimensione del record R in byte.
- Calcolare il fattore di blocco bfr e il numero di blocchi b del file supponendo un'organizzazione unspanned
- Si supponga che il file sia *ordinato* con il campo chiave CF e che si voglia creare un *indice primario* su CF. Calcolare:
 - il fattore di blocco dell'indice $bfri$
 - il numero delle voci e il numero di blocchi dell'indice primario;
 - il numero di livelli necessari se si vuole passare a un indice multilivello;
 - il numero totale di blocchi richiesti dall'indice multilivello;
 - il numero di accessi necessari per cercare e reperire un record dal file, dato il valore di CF, usando l'indice primario;
 - il numero di accessi necessari per cercare e reperire un record dal file, dato il valore di CF, usando l'indice multilivello;
- Si supponga che il file *non sia ordinato* tramite la chiave CF e si voglia creare un *indice secondario* su CF. Ripetere l'esercizio precedente (c) per l'indice secondario e confrontarlo con l'indice primario.

Esercizio -1

Si consideri un disco con dimensione di blocco $B = 512$ byte. Un puntatore a blocco è lungo $P = 6$ byte. Il file ha $r = 30.000$ record della tabella IMPIEGATO di *lunghezza fissa*. Ciascun record ha i seguenti campi:

- NOME (30 byte),
- CF (16 byte),
- NUMERO_DIPARTIMENTO (9 byte),
- TELEFONO (9 byte),
- DATA_NASCITA (8 byte),
- CODICE_LAVORO (4 byte).
- Un ulteriore byte viene utilizzato come indicatore di cancellazione.

a) Calcolare la dimensione del record R in byte.

77 byte

b) Calcolare il fattore di blocco bfr e il numero di blocchi b del file supponendo un'organizzazione unspanned

$$Bfr = \lfloor 512/77 \rfloor = 6$$

$$\text{numero blocchi} = \lceil 30.000/6 \rceil = 5000$$

Esercizio -1

$B = 512$ byte

$r = 30.000$ record

Fattore di blocco del record =6

$P = 6$ byte

Dimensione del record R in byte= 77 byte

numero blocchi file di dati=5000

c) Si supponga che il file sia *ordinato* con il campo chiave CF (16 byte) e che si voglia creare un *indice primario* su CF. Calcolare:

i. il fattore di blocco dell'indice *bfri*

Dimensione voce= Dimensione CF + dimensione puntatore= 16+6= 22

Fattore di blocco dell'indice $bfri = \lfloor 512/22 \rfloor = 23$

ii. il numero delle voci e il numero di blocchi dell'indice primario

Il file è ordinato sul campo CF (campo dell'indice primario). E' necessario una voce per ogni blocco dati (indice sparso)

Numero voci indice primario= 5000

Numero blocchi dell'indice primario: $\lceil 5.000/23 \rceil = 218$

Esercizio-1

$B = 512$ byte

$r = 30.000$ record

Fattore di blocco del record =6

fattore di blocco dell'indice =23

$P = 6$ byte

Dimensione del record R in byte= 77 byte

numero blocchi file di dati=5000

numero blocchi dell'indice primario:= 218

c) Si supponga che il file sia *ordinato* con il campo chiave CF (16 byte) e che si voglia creare un *indice primario* su CF. Calcolare:

iii. il numero di livelli necessari se si vuole passare a un indice multilivello;

2° livello: indice per un file ordinato (indice primario)

numero voci secondo livello:

218 (una per ogni blocco dell'indice primario – indice sparso)

numero blocchi per secondo livello:

$$\lceil 218/23 \rceil = 10$$

serve un altro livello?

Si, per indicizzare i 218 voci nei 10 blocchi

3° livello: indice per un file ordinato (anche il secondo livello è ordinato)

numero voci terzo livello:

10 (una per ogni blocco dell'indice di secondo livello – indice sparso)

numero blocchi per terzo livello:

$$\lceil 10/23 \rceil = 1$$

serve un altro livello?

No. 10 voci si indicizzano con un blocco solo

iv. il numero totale di blocchi richiesti dall'indice multilivello;

$$\text{Blocchi 1° livello} + \text{blocchi 2° livello} + \text{blocchi 3° livello} = 218 + 10 + 1 = 229$$

v. il numero di accessi necessari per cercare e reperire un record dal file, dato il valore di CF, usando l'indice primario;

$$\text{ricerca binaria su indice primario} + \text{accesso al blocco dati} \Rightarrow : (\log_2 218) + 1 \Rightarrow 8 + 1 = 9$$

vi. il numero di accessi necessari per cercare e reperire un record dal file, dato il valore di CF, usando l'indice multilivello;

$$\text{uno accesso al blocco per livello} + \text{tempo accesso al blocco dati} \Rightarrow 3 + 1 = 4$$

Esercizio -1

$B = 512$ byte

$r = 30.000$ record

Fattore di blocco del record =6

fattore di blocco dell'indice =23

$P = 6$ byte

Dimensione del record R in byte= 77 byte

numero blocchi file di dati=5000

numero blocchi dell'indice primario:= 218

d) Si supponga che il file **non sia ordinato tramite** la chiave CF e si voglia creare un *indice secondario* su CF. Ripetere l'esercizio precedente (c) per l'indice secondario e confrontarlo con l'indice primario.

i. il fattore di blocco dell'indice *bfri* -- non cambia

ii. Indice secondario: il numero delle voci e il numero di blocchi dell'indice secondario

Il file NON è ordinato sul campo CF. E' necessario una voce per ogni record (indice denso)

Numero voci indice secondario= 30000

Numero blocchi dell'indice secondario: $\lceil 30.000/23 \rceil = 1305$

iii. il numero di livelli necessari se si vuole passare a un indice multilivello;

2° livello: indice per un file ordinato (indice secondario è ordinato)

numero voci secondo livello: 1305 (una per blocco dell'indice secondario – indice sparso)

numero blocchi per secondo livello: $\lceil 1305/23 \rceil = 57$

serve un altro livello?

Si, per indicizzare i 1305 voci nei 57 blocchi

3° livello: indice per un file ordinato (anche il secondo livello è ordinato)

numero voci terzo livello: 57 (una per blocco per l'indice di secondo livello – indice sparso)

numero blocchi per terzo livello: $\lceil 57/23 \rceil = 3$

serve un altro livello?

Si. 57 voci richiedono 3 blocchi

4° livello: indice per un file ordinato

numero voci terzo livello: 3 (una per blocco per l'indice di secondo livello – indice sparso)

numero blocchi per terzo livello: $\lceil 3/23 \rceil = 1$

serve un altro livello?

No. 3 voci richiedono 1 blocco solo

iv. il numero totale di blocchi richiesti dall'indice multilivello;

Blocchi 1° livello + blocchi 2° livello + blocchi 3° livello + blocchi 4° livello = $1305 + 57 + 3 + 1 = 1366$

v. il numero di accessi necessari per cercare e reperire un record dal file, dato il valore di CF, usando l'indice secondario;

ricerca binaria su indice primario + accesso al blocco dati $\Rightarrow (\log_2 1305) + 1 = 11 + 1 = 12$

vi. il numero di accessi necessari per cercare e reperire un record dal file, dato il valore di CF, usando l'indice multilivello;

uno accesso al blocco per livello + accesso al blocco dati: $4 + 1 = 5$