

Esercizi: EX 1. STRUTTURE PRIMARIE

Si consideri una tabella $R(A,B)$, con A di 18 byte e B di lunghezza variabile, massimo 20 byte. R è memorizzato in un file disordinato, con blocchi dati di $B=400$ byte e con record, di lunghezza variabile, che richiedono 2 byte per la gestione (e.g., cancellazione, puntatori etc). Gli spazi dei record eliminati sono marcati come liberi e non riutilizzati. Nel caso di modifica del record, dato la lunghezza variabile, se la lunghezza del record cambia, l'operazione viene gestita come cancellazione ed inserimento del record modificato. La tabella è inizialmente vuota

Si supponga di eseguire le seguenti operazioni:

1. Inserimento di 1000 record con campo B di lunghezza nulla
2. Aggiornamento di 100 record, con modifica del valore di B, che occupa ora 20 byte
3. Riorganizzazione del file con ricompattazione dei blocchi

Indicare il numero di blocchi occupati dalla tabella dopo ciascuna operazione.

Esercizi: Ex.2 INDICI PRIMARI/SECONDARI/MULTILIVELLO

Si consideri la seguente configurazione:

Dimensione blocco, $B = 512$ byte.

Puntatore a blocco, $P = 6$ byte

#record $r = 100.000$ IMPIEGATO

Campi record IMPIEGATO :

NOME (30 byte),

SSN (9 byte),

NUMERO_DIPARTIMENTO (9 byte),

DATA_NASCITA (8 byte),

SESSO (1 byte),

Un ulteriore byte viene utilizzato come indicatore di cancellazione.

D1) Si assuma che il file dati di IMPIEGATO sia ordinato con il campo chiave SSN e che si è creato un indice primario su SSN. Calcolare:

- il fattore di blocco dell'indice bfri

- il numero delle voci e il numero di blocchi dell'indice del primo livello;
- 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 SSN, usando l'indice primario
- il numero di accessi necessari per cercare e reperire un record dal file, dato il valore di SSN, usando l'indice multilivello

D2. Si supponga che il file non sia ordinato, e si voglia creare un indice secondario su SSN.

- il fattore di blocco dell'indice bfri
- il numero delle voci e il numero di blocchi dell'indice secondario;
- 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 SSN, usando l'indice secondario
- il numero di accessi necessari per cercare e reperire un record dal file, dato il valore di SSN, usando l'indice multilivello

Si considera una tabella $R(A, C, \dots)$ con 1.000.000 di record di lunghezza fissa di 100 byte. In R , A è campo chiave. La dimensione di A è 4 byte, la dimensione di C è 4 byte. La memoria secondaria è organizzata in blocchi di dimensione $B = 1000$ byte.

Considerare il seguente carico di lavoro:

- Op1. inserimento di un record (con verifica del vincolo di chiave), con frequenza giornaliera $f_1 = 1$;
- Op2. ricerca di un record sulla base del valore della chiave A , con frequenza giornaliera $f_2 = 1000$;
- Op3. ricerca di un record sulla base del valore dell'attributo C , con frequenza giornaliera $f_3 = 10.000$;

Assumere, inoltre, che di notte il sistema esegue la riorganizzazione delle strutture fisiche

Definire i costi giornalieri valutando le seguenti soluzioni

- (A) Struttura primaria ordinata sull'attributo A , con eventuali inserimenti effettuati in coda al file
- (B) Struttura primaria non ordinata con indice B^+ -tree sull'attributo A
- (C) Struttura primaria ordinata sull'attributo A con indice B^+ -tree su campo C , con eventuali inserimenti effettuati in coda al file
- (D) Struttura primaria non ordinata, con indice B^+ -tree sull'attributo A e indice B^+ -tree sull'attributo C

Si assuma, indici B^+ -tree sono usati con fattore di riempimento a 70%, un puntatore P al blocco occupa 6 byte, un puntatore P_r al record occupa 7 byte. Come costo stimare il tempo di esecuzione in termini di accessi alla memoria secondaria.