

Due relazioni:

- Attributi e riferimenti hanno la stessa taglia
- $R(A,B,C)$ 700 pagine, ognuna con 80 tuple. L'attributo A assume 560 valori distinti uniformemente distribuiti su tutte le tuple
- $S(\underline{D},\underline{E},F,G,H,L)$ ha 1000 pagine, e un indice B+ su $\langle D,E \rangle$

Domanda 1. Quanto costa in termini di accesso a pagine la seguente query *'select * from R where A = 20'* nel caso in cui R sia un heap file, un sorted file e un hashed file? quante tuple saranno emesse nel risultato?

Domanda 2. Quante tuple sono contenute per pagina nella relazione S?

Domanda 3. Per ognuna delle possibili organizzazioni di S (head, sorted, hashed):

1. che alternative (1,2,3) possono essere utilizzate per l'implementazione dell'indice?
2. In quali casi l'indice può essere clustered o unclustered?
3. In quali casi l'indice può essere denso o sparso?
4. Nel caso di indice denso, quanti livelli ha l'indice?

Domanda 4. Si assuma di avere una memoria principale con 4 frame, ognuno in grado di ospitare una pagina. Considerando la seguente query *'select * from R,F where A = 20 and B=D and C=E'*, che strategia utilizzare per calcolare la suddetta query? Quale sarebbe il costo per le differenti organizzazioni delle relazioni R e S (heap, sorted, hash).