

Ordinamento (DBMS)

Scopi: implementazione ORDER BY, JOIN, DISTINCT, GROUP BY.

Avendo a disposizione b buffer, si può ordinare con un merge sort a $z = b - 1$ vie:

1. si caricano in memoria b pagine alla volta, si ordinano con quicksort e si scrivono su disco, in modo da ottenere $\lceil p/b \rceil$ run ordinati;
2. si fondono fino a z (il b -esimo buffer serve per l'output) run (di lunghezza l) per ottenere run di lunghezza zl , ripetendo finché tutti i dati non sono ordinati.

Con p pagine, il costo in numero di letture/scritture di pagine è $2p(1 + \lceil \log_{b-1}(p/b) \rceil)$: $2p$ per l'ordinamento interno e $2p$ per ciascuno dei $\lceil \log_{b-1}(p/b) \rceil$ passi di merge (o $\lceil \log_{b-1} p \rceil$ se si ordinano singolarmente le pagine nel passo 1).