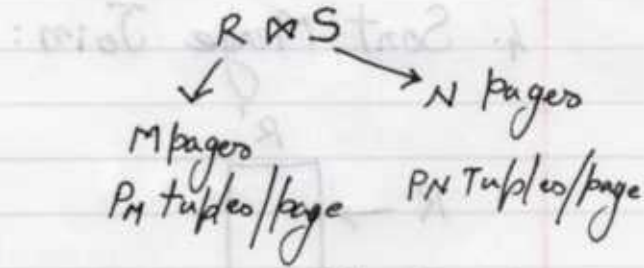


30/3/09

Database Management Systems

JOIN PROCESSING

1. Simple Nested Loop Join



For each tuple $r \in R$ — outer relation
 For each tuple $s \in S$ — inner relation
 if (r and s is a match)
 output r, s

Focus: Equality Search.

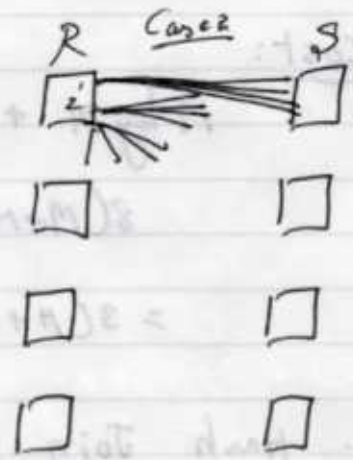
$$M + \underbrace{P_M \cdot M \cdot N}_{\text{total \# of tuples in } M}$$

$$M + M \cdot N$$

An obvious improvement is to find the matches page by page, I/O complexity is

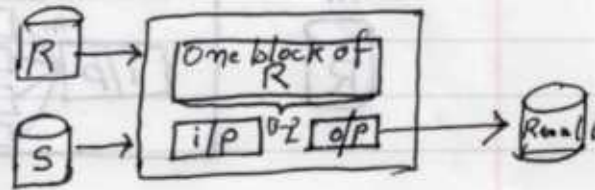
uses 3 pages.

buffer



2. Block-Based Nested Loop Join: $B \gg 3$ pages.

$$M + \frac{M}{B-2} \cdot N$$



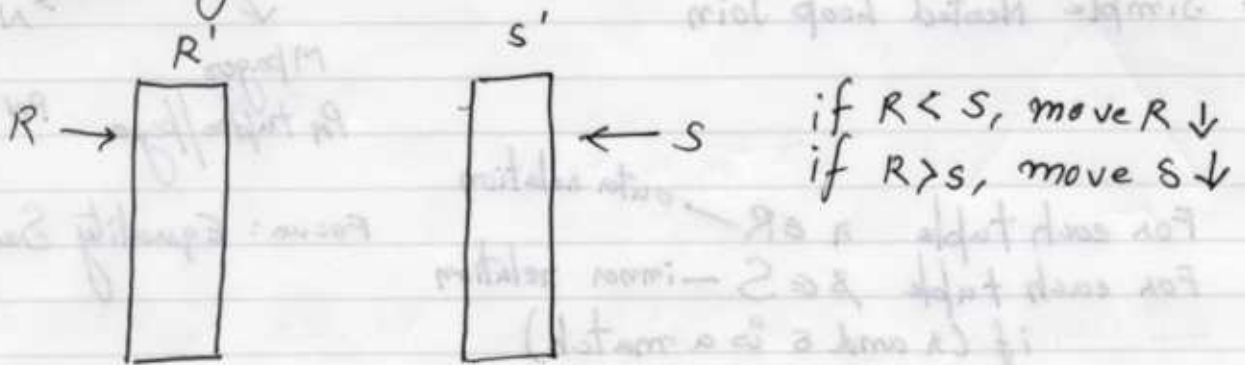
3. Index Based Nested Loop Join:

For $r \in R$

for each page P in R ,
 goto indexed table S ,
 and find all matches to tuples in P .

$$M + C \cdot M \cdot P_m$$

4. Sort-Merge Join:



Cost:

$$M \log M + N \log N + M + N$$

$$8(M+N) - (MN)$$

$$= 3(M+N) \text{ in practise}$$

this is the average case, the worst case is when all tuples in R match all tuples in S - this part becomes MN. However, this may not happen that often.

5. Hash Join:

