Name 1: Date:
Name 2:

Assume:

- Relation R(a,b) contains 10,000 tuples, and has 10 tuples per block
- Relation S(a,c) contains 2,000 tuples and has 10 tuples per block
- Both relations are stored as simple heap files
- Neither relation has any indexes built on them

Part A:
Compute SELECT * FROM R NATURAL JOIN S;

Assume that the join is computed using a block oriented, simple nested loop. What is the minimum number of memory buffer blocks required to do this join in exactly one pass?

Part B:

How many memory blocks do we need to calculate $\Pi_a R \cap \Pi_b S$ in one pass?

Part C:
How many block reads do we need to compute the following query in one pass. Assume a naive evaluation plan (execute the intersection first, then the join)

```
SELECT a FROM
  R JOIN
    (SELECT b FROM R
     INTERSECT
     SELECT c FROM S);
```

Part A:

To do the join in one pass we need to have enough memory for the smallest table + 1 block.

$B(R) = 10^5/10 = 10^4$ blocks

$B(S) = \frac{2 \cdot 10^3}{10} = 2 \cdot 10^2$ blocks.

$B(S) < B(R)$

Read S. Read 1 block of R.

$\Rightarrow$ We need $B(S) + 1$ block memory.

$$\text{Cost} = B(S) + \left\lceil \frac{B(R) \cdot B(S)}{B(S)} \right\rceil$$

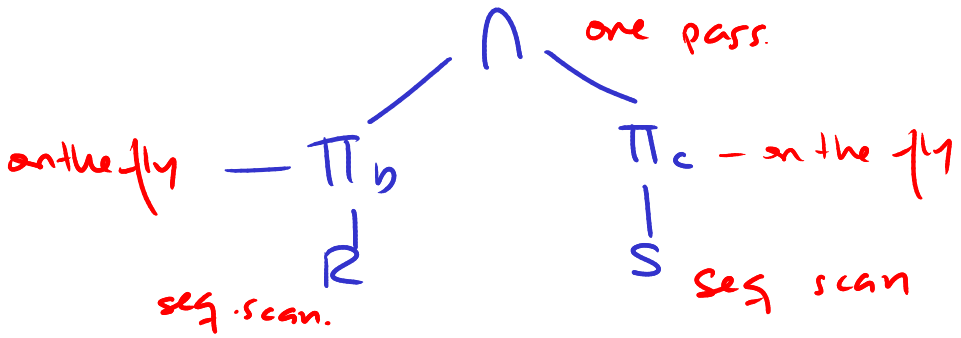
↑ memory for the join

$$= B(S) + B(R)$$

One pass!!

Part B:

To do this join we:



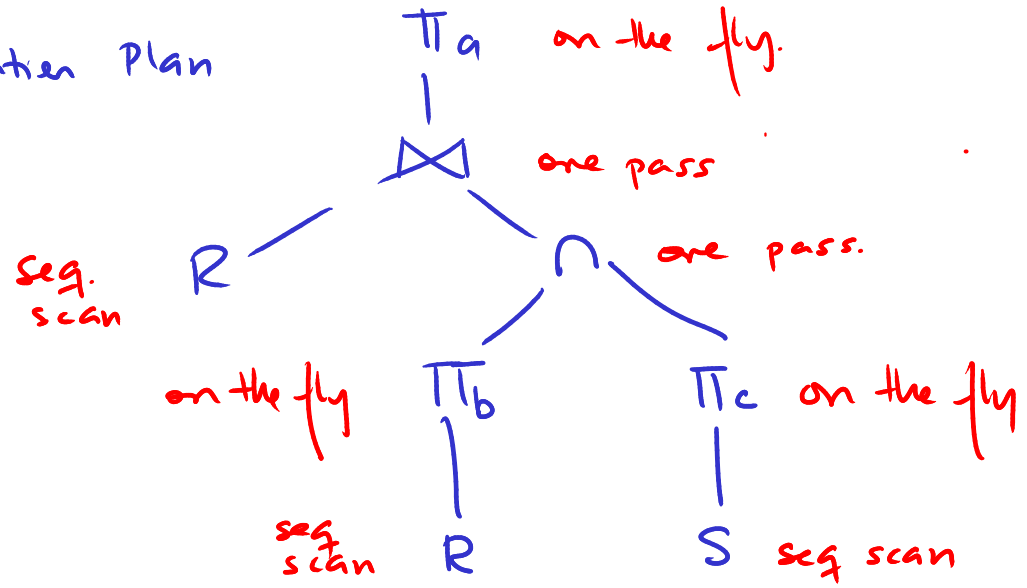
To do it in one pass we need enough memory to either place $\Pi_b R$ in memory or $\Pi_c S$ in memory. We don't know anything else about either attribute.

Memory:

$$M \geq \min(B(\Pi_b R), B(\Pi_c S)) + 1 \text{ block to read } R + 1 \text{ block to read } S$$

Part C

Evaluation Plan



From part B we know we need enough memory for either $\Pi_c S$ or $\Pi_b R$. (Whichever is smaller.)

But we also need memory for the join, so the join can start as soon as the $\cap$ generates tuples.

Total memory required >=

Memory Intersect + Memory Join

The memory for the join is the smaller of $B(R)$ and $B(\Pi_c S \cap \Pi_b R)$. Given it is an intersect with a subset of R:

$$B(\Pi_c S \cap \Pi_b R) < B(R)$$

To read R.

$\Rightarrow$ Memory Join is $B(\Pi_c S \cap \Pi_b R) + 1$.

So Total Memory $\geq \min(B(\Pi_c S), B(\Pi_b R)) + 2$

$+ B(\Pi_c S \cap \Pi_b R) + 1$.