

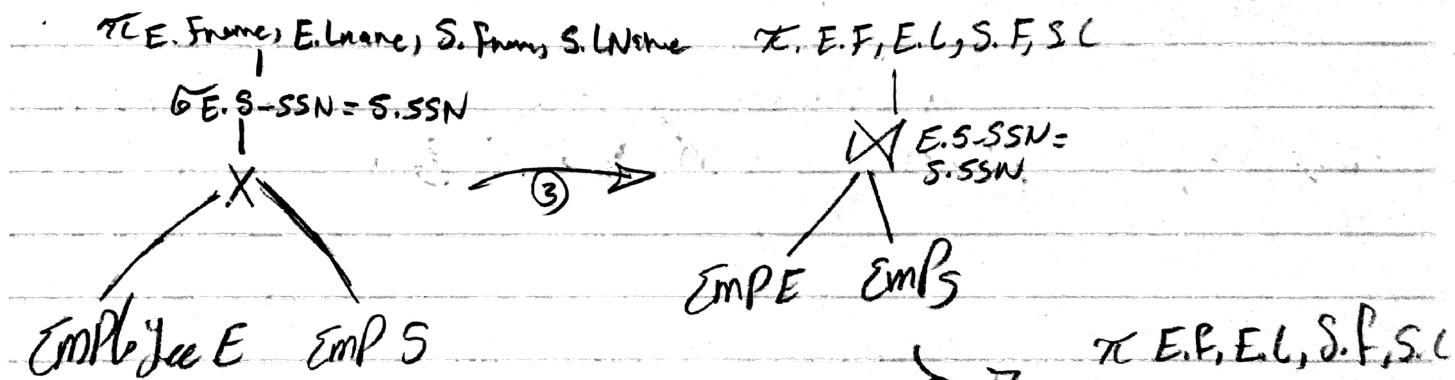
ADB

Q.1

for each query  
Sheet 3  
↳ Query tree  
↳ Initial Q tree  
↳ Compare & select

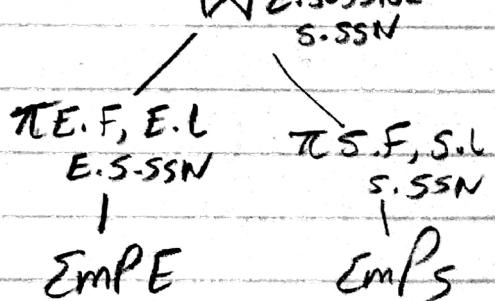
Query: Select E.FName, E.LName, S.FName, S.LName  
FROM EmployeeE, EmployeeS  
Where E.SuperSSN = S.SSN

Initial Tree



Can

- ① Move Selection Down? No
- ② Apply most restrictive select? No
- ③ Replace [Cartesian Product] and [Select] with join? Yes
- ④ Move Projection down? Yes



الأول بالtree إذا لم ي Reduce  
وأول الأسئلة دوافع

يمكن إيجاد ذلك بـ  $\exists$  ثم حاجة  
التي فيه  $\exists$  للjoin

Optimized Tree

$\pi \Rightarrow Proj(Select) \rightarrow Table.no$

$G \Rightarrow Selection(Where) \rightarrow Reduce$

$X \Rightarrow Cartes. Prod att.no$

$M \Rightarrow join on Common Col C$

$\Rightarrow X \rightarrow G_C = M_C$

Recall:

$\pi_{SSN \in Emp}$   
 $\sigma_{ESSN=workson}$   
 $PNO \in workson$   
 $PNAME \in Proj$   
 $\sigma_{PName=GProj}$   
 $\pi_{PNAME=GProj}$   
 $\pi_{PNAME="ProjX"}$

Follow Q. 1

Query 27: Select FName, LName, I.I.Salary  
 FROM Employee, workson, Project  
 WHERE SSN=ESSN AND PNO=PNum AND  
 $\pi_{FName, LName, I.I.Salary}$        $\pi_{FName, LName, I.I.Sal}$

Initial Query

$\sigma_{SSN=ESSN}$   
 $\sigma_{PNO=PNum}$   
 $\sigma_{PName="X"}$

$\Rightarrow$

$\sigma_{SSN=ESSN}$

X

EMP

workson Proj

$\sigma_{PNO=PNum}$   
 $\sigma_{PName="ProjX"}$   
 $\sigma_{Proj}$

③

$\pi_{FName, LName, I.I.Sal}$   
 $\sigma_{SSN=ESSN}$

- ① Move Select down? Yes
- ② Apply Nest Restrictive join? No
- ③ Replace  $[X \rightarrow \sigma]$  with  $[G]$ ? Yes
- ④ Move Proj down? Yes

$\sigma_{PNO=PNum}$  EMP

$\sigma_{PName="ProjX"}$  Proj

$\pi_{FName, LName, I.I.Sal}$

$\sigma_{SSN=ESSN}$

$\sigma_{PNO=PNum}$

$\pi_{FName, LName, Sal, SSN}$

EMP

$\pi_{PNO, ESSN}$

$\pi_{PNum}$

$\sigma_{PName="ProjX"}$

$\pi_{PName, PNum}$

Proj

15.14)  $N_{\text{blocks}} = 4096$  ] How many passes will be needed in the external merge sort?  
 $N_{\text{buffers}} = 64$

$$* I/Os = 2N \left( 1 + \lceil \log_{B-1} \lceil N/B \rceil \rceil \right)$$

↑  
Passes      ↓  
Passes in N

$$* \text{noPasses} = 1 + \lceil \log_{63} \lceil \frac{4096}{64} \rceil \rceil = 1 + \lceil 1.0038 \rceil = 1 + 2 = 3$$

15.15 → مُضافة  
الآخر

15.17

- \* End of file, we can't do all the possible aggregate func, for instance we can do the [MIN] func, since the index points to the smallest key value in the data block. we can apply the [MAX] func if the index was ordered descendingly!
- \* we can apply the [Count] func if we are using secondary from key
- \* we can't do the [SUM] nor the [AUG] functions.

15.21

left outer join



وكان C.C كـ join-conditions لو اراد left outer join

فيه حاجة في R تـي  $v_i$  في S تـي  $s_j$  مـع  $v_i < s_j$  ولو

مـعنى اـ جـع  $v_i$  وـ عـبر عن  $\langle v_i, \text{Null} \rangle$  بـ رـاـضـتـك

Col C كـ  $s_j$  و Col C كـ R كـ join  $\xrightarrow{\text{sort-merge}}$

فـ يـ جـع اـ خـصـوـة دـيـ عـادـي

بعد كـ دـيـ يـ جـع اـ خـصـوـة دـيـ عـادـي

لـمـعـ ماـ يـ لـاـتـيـ حـاجـةـ دـيـ شـامـلـيـ مـنـيـ  $v_i$  يـ جـعـيـ

لـمـعـ ماـ يـ لـاـتـيـ حـاجـةـ دـيـ شـامـلـيـ مـنـيـ  $v_i$  يـ جـعـيـ

step1  $\Rightarrow$  sort+RDS on Col C  
 step2  $\Rightarrow$  while( $s_i < r_j$ )  $i++$   
اللى هـيـ جـعـيـ  
 if  $s_i == r_j$  return  $\langle r_j, s_i \rangle$   
 if  $s_i > r_j$  return  $\langle r_j, \text{Null} \rangle$

الشرح في  
الى بعدها

15.22] Comp the Cost of 2-diff Q-Plans for this Qwry ,  
 [ Salary > 40K , Select (EMP X DNO = DNumber Department )  
 $\rightarrow$  ~~Salary > 40K (EMP D)~~ Given E has 2K blocks

Assume  
min w/o  
dups

$\frac{500 \text{ uniqueSal}}{\text{Dnum}} + \frac{\text{no of leaves} \in [B + \text{tree has } (1 + \text{leaf}) \text{ levels}]}{\text{Dnum}} + \text{with 10K records}$   
 The Salary = 50  $\rightarrow$  D has 5 blocks with 50 records .  
 Salary  $\in [1, 500K]$

### Plan 1

- \* % of Salary Chosen =  $\frac{500 - 400}{500} = \frac{1}{5}$  Assume 100 value  $\rightarrow 40K$
- \* No of Salaries include block access = 1 block +  $\frac{1}{5} \times \text{No of leaves} = 1 + \frac{1}{5} \times 50 = 11$
- \* No of records with Salary  $> 400 = \frac{1}{5} \times \text{total no of records} = \frac{1}{5} \times 10000 = 2000$  emp records
- \* with uniform access, no of block access  $\leftarrow$  Assume, Records 10000 Needed, to read records after index = 2000 : ~~blocks cost~~  $\rightarrow$  ~~blocks cost~~  $\rightarrow$  ~~blocks cost~~
- \* To store this result we need = 400  $\rightarrow$  2K  $\in$  10K  $\rightarrow$   $\frac{400}{2K} = 200$   $\rightarrow$   $\frac{200}{10K} = 20$   $\rightarrow$   $20 \times 2K = 400$
- \* Total Cost of Selection =  $11 + 2000 + 400$   $\rightarrow \frac{400}{5} = 80$

\* Cost of join =  $(\text{Dep}) \text{blocks} + (\text{Dep}) \text{blocks} * (\text{Temp}) \text{blocks} + \text{cost}$   
 $= 5 + 5 * 400$  Plan 2  $\rightarrow$   $\frac{400}{5} = 80$

\* Total Cost = Cost of Selection + Cost of Join  
 $= (11 + 2000 + 400) + (5 + 5 * 400) = 4416$  block ACCESS

Note: if we have 5 enough memory buffer pages available during the join, we could store all 5 blocks of the dep table there. Hence, join cost =  $5 + 400 = 405$

$\therefore$  Total Cost =  $(11 + 2000 + 400) + 405 = 2816$

Follow 15.22

Plan 2

\* Selection Cost = Same as Plan 1

- \* Join Cost  $\rightarrow$  nested loop we will read all the records  
"emp"

We can use Pipelining

to reduce the writing overhead

لويات كم (احسب) في بلاك (1) في write over head

2

\* هنا عندنا بلاك 1 و 2 ، فنحجب نعمل Selection قبل ال Join و بعده و نقارب ال cost بعدي ! خذ بالكل لعمان نفس ال Select تعميرياً

\* عاشر بلاك 5d  $\rightarrow$  assumption  $\rightarrow$  يقول دات 1/5 من ال EMP يتحققوا الشرط ! دعف من ال 15K موقف  $\rightarrow$  هطلع 2K جس !

\* عناصر من عارق ال block 5d ، تكونت لو ال 10K ادخرت في 2K من ال blocks يبقى 2K ليخزنوا في 400 ، الى هي الخمس بالصفحة كده عددى 400 بلوك بينها وروا عى 2K بلوك .

\* ال Index بعض ؟ هو تاسع عنابر فيكون  $1 + \text{عدد ال leaves}$  هو تاسع ؟ ال Index بيتبعي على ال Dep  $\rightarrow$  ال الي فتحوا 50 Years  $\rightarrow$  فقط  $\rightarrow$  assumption EMP متوزع عيس بشكل uniform بالواسطي

ال خاتمة ال Index هو  $11 = 15 * 50 + 1$  ، الى هو خصم برمونو .

\* كده عنابر 10 عمل  $\rightarrow$  Select cost =  $11 + 2000 + 400$   $\rightarrow$  Block access

\* بالنتيجة لا Join حسب المعايير ، لو  $M + MN = 5$   $\rightarrow$  simple nested join يبيح ادا M هناد المغير ، الى هو ال Dep بـ 5.

\* في بلاك 10 كان لـ N بـ 400 و بلوك 5 ال N بـ 2000  $\rightarrow$  ادا Join عناه في بلاك 5 معملاستي Select قبل ال Dep

6 Salary > 400

$M_{DN} = DN_{num}$

Emp Dep

**15.15** APPROX Cost func for: Proj, Union, Intersection  
 Set Diff, Cartesian Prod  
 Assume Relation R, S  
 with blocks M, N respect  
 ALGos

- \* Projection on R:
  - No duplicates  $\Rightarrow I/Os = M + M \quad \begin{matrix} \text{read and select col} \\ \text{write the} \\ \text{blocks} \end{matrix}$
  - Duplicates  $\Rightarrow I/Os = 2M(1 + \lceil \log_2 M \rceil) \quad \begin{matrix} \text{2-way} \\ \text{external} \\ \text{sort} \end{matrix}$   
 "Not accurate at all" eliminate duplicates from the col ↴

- \* Set Operation:
  - use sort-merge join  $\Rightarrow I/Os = \left[ \begin{matrix} 2M(1 + \lceil \log_2 M \rceil) \\ 2N(1 + \lceil \log_2 N \rceil) \end{matrix} \right] + \left[ \begin{matrix} M \\ N \end{matrix} \right]$   
 Sort Merge

- \* Cartesian Prod.
  - simple nested loop  $\Rightarrow M + MN \quad \text{assume } M < N$
  - Block nested loop  $\Rightarrow M + \lceil \frac{M}{B-2} \rceil N$
  - Sort-Merge ---