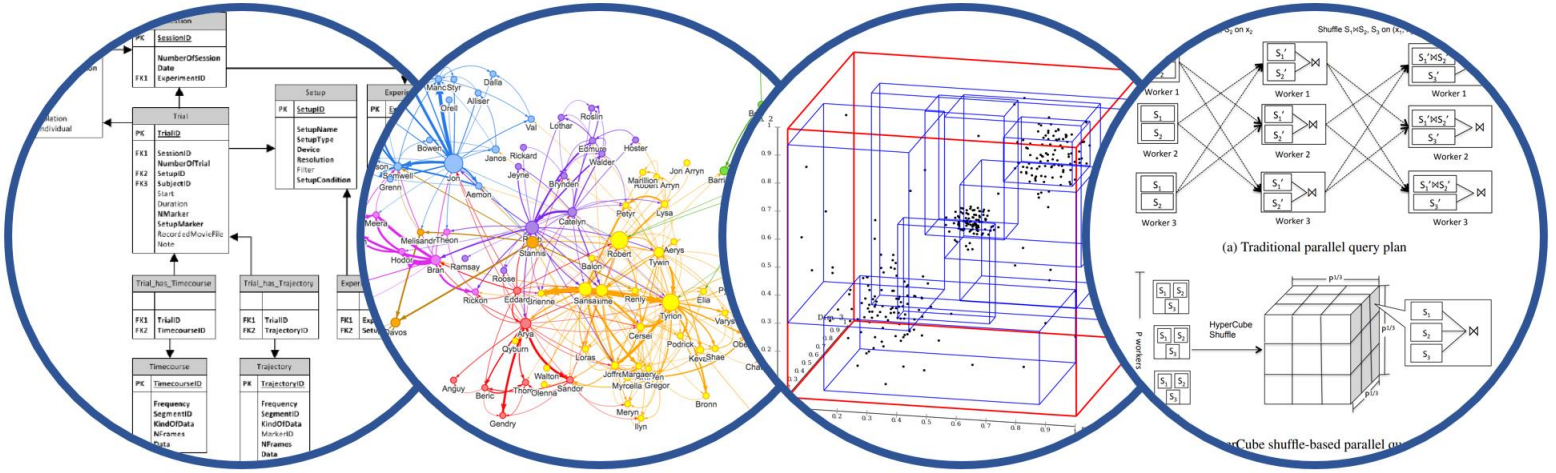


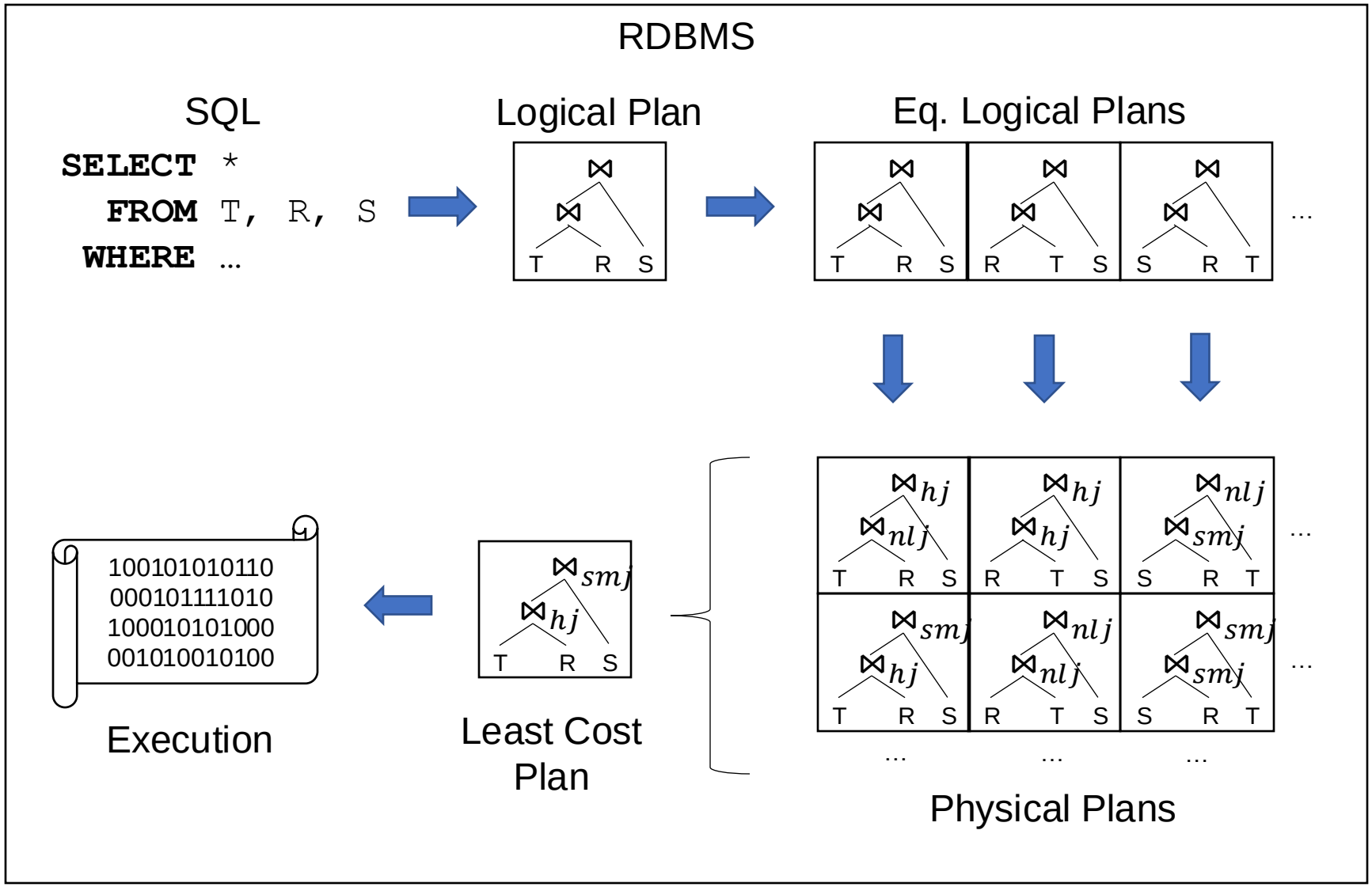
Introduction to Data Management

Query Optimization

Paul G. Allen School of Computer Science and Engineering
University of Washington, Seattle

Query Optimization

Plan Enumeration



Query Optimization

- Space space
- Cardinality and cost estimation
- Search algorithm

Search Space

Search Space

- Optimizer explores alternative plans P_1, P_2, …
- The set of all plans is called the search space
- Two parts:
  - Rewrite rules
  - Access path selection

Rewrite Rules

Rewrite Rules

- A rewrite rule is an identity in relational algebra
- Similar to identities in algebra, e.g.

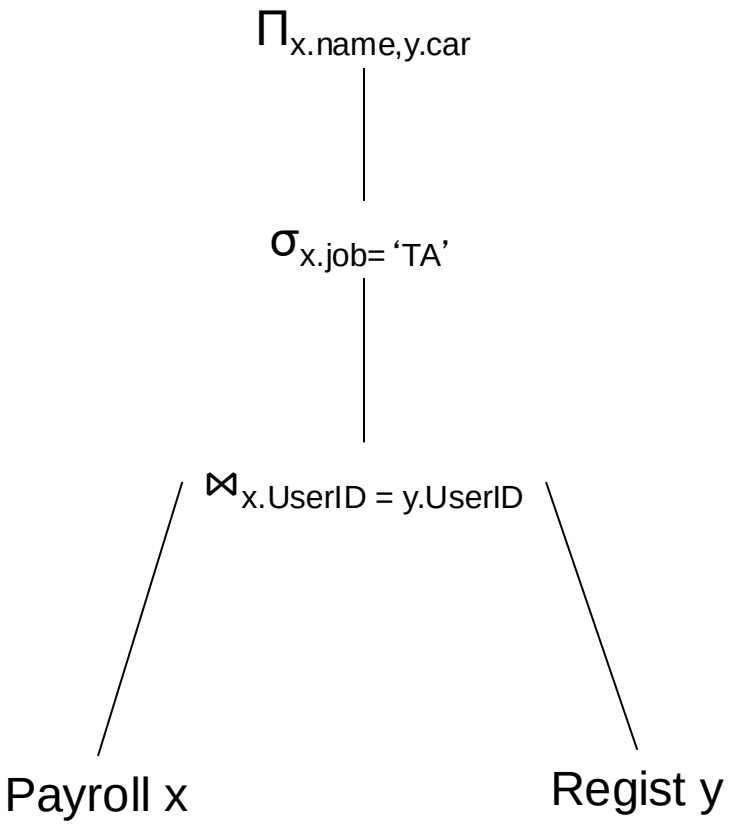
```
a+b=b+a                 // commutativity
a*(b+c)=a*b+a*c         // distributivity
…
```

- An optimizer applies a set of rewrite rules

Rewrite Rules: Predicate Pushdown

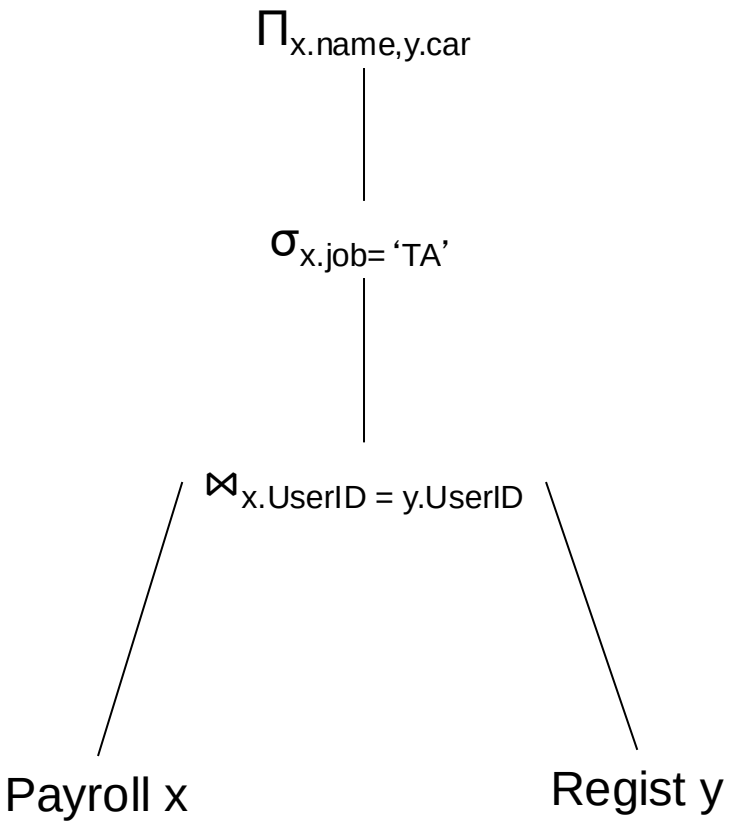
```
SELECT x.name,y.car
FROM Payroll x, Regist y
WHERE x.UserID = y.UserID
 and x.job='TA';
```

Rewrite Rules: Predicate Pushdown

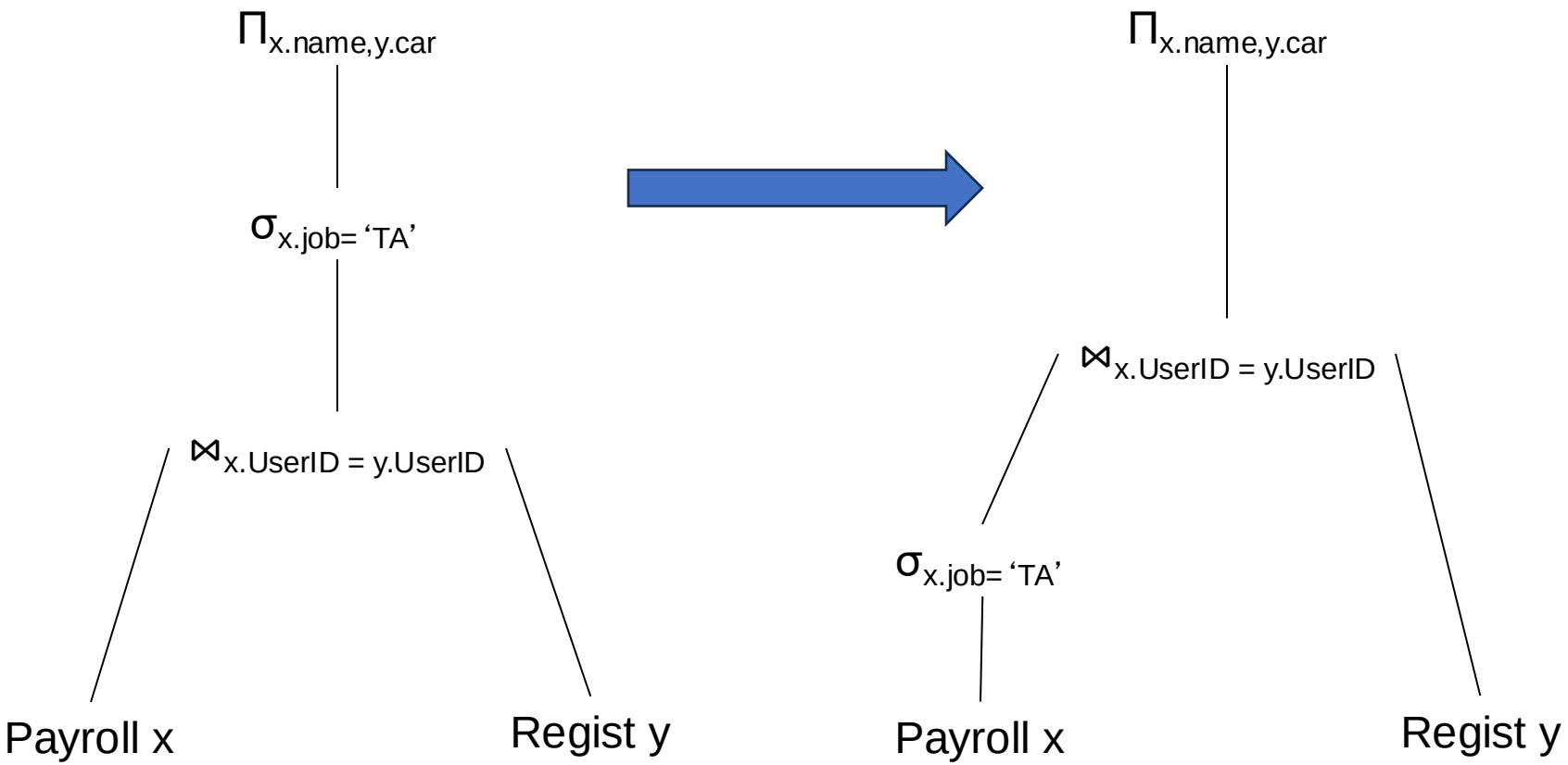



```
SELECT x.name,y.car
FROM Payroll x, Regist y
WHERE x.UserID = y.UserID
 and x.job='TA';
```

Rewrite Rules: Predicate Pushdown

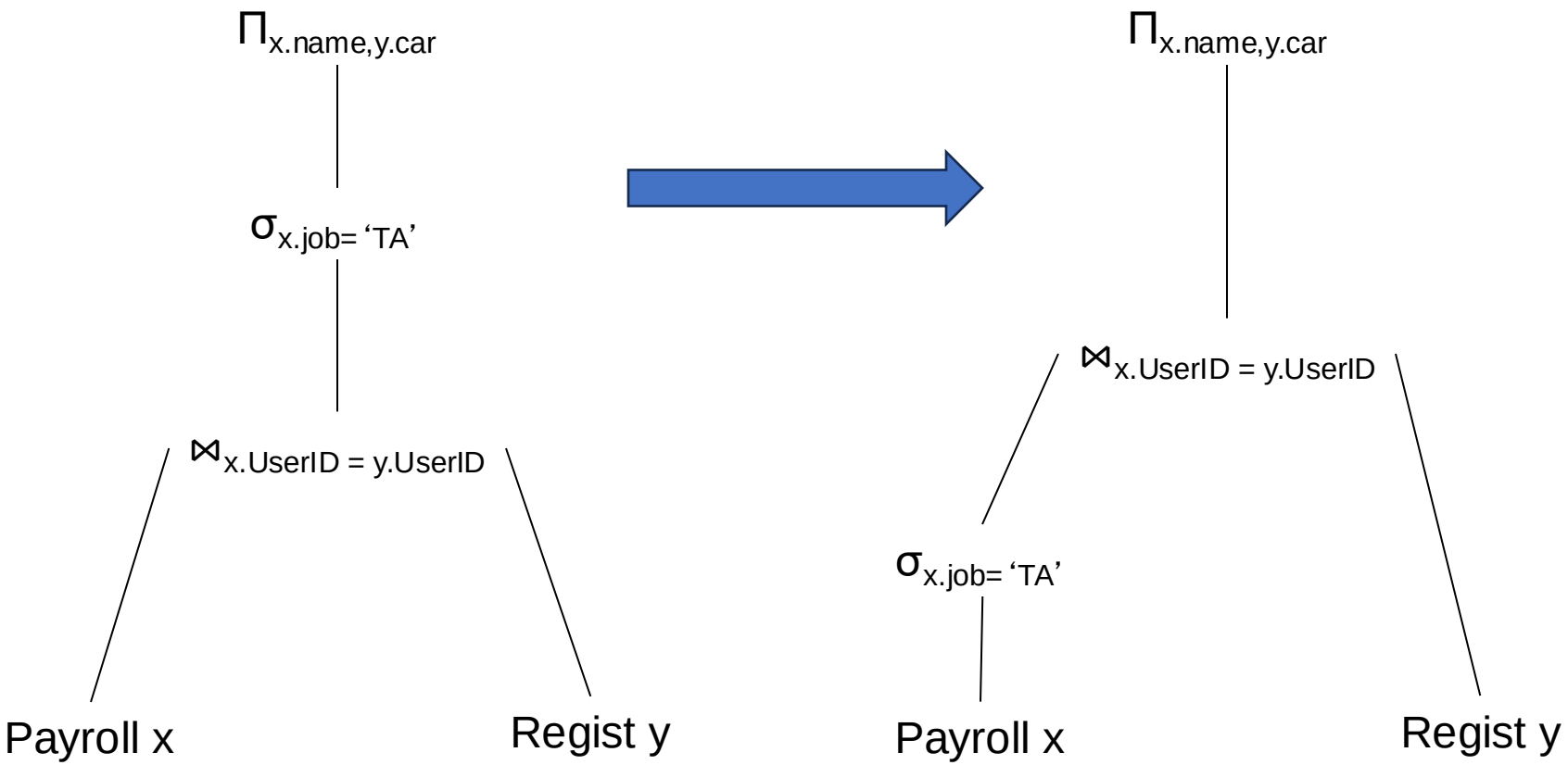


Rewrite Rules: Predicate Pushdown

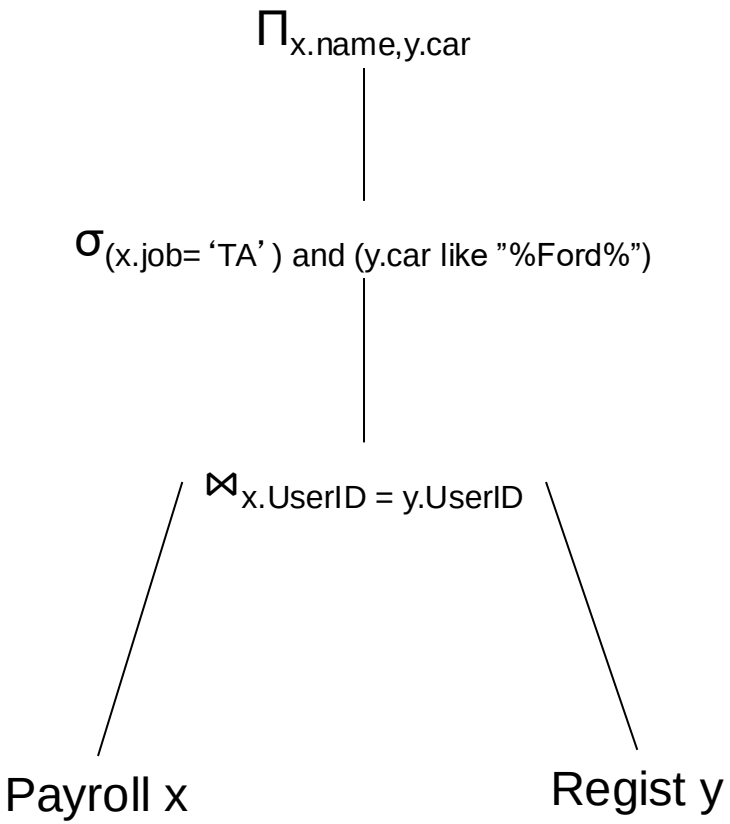


Rewrite Rules: Predicate Pushdown

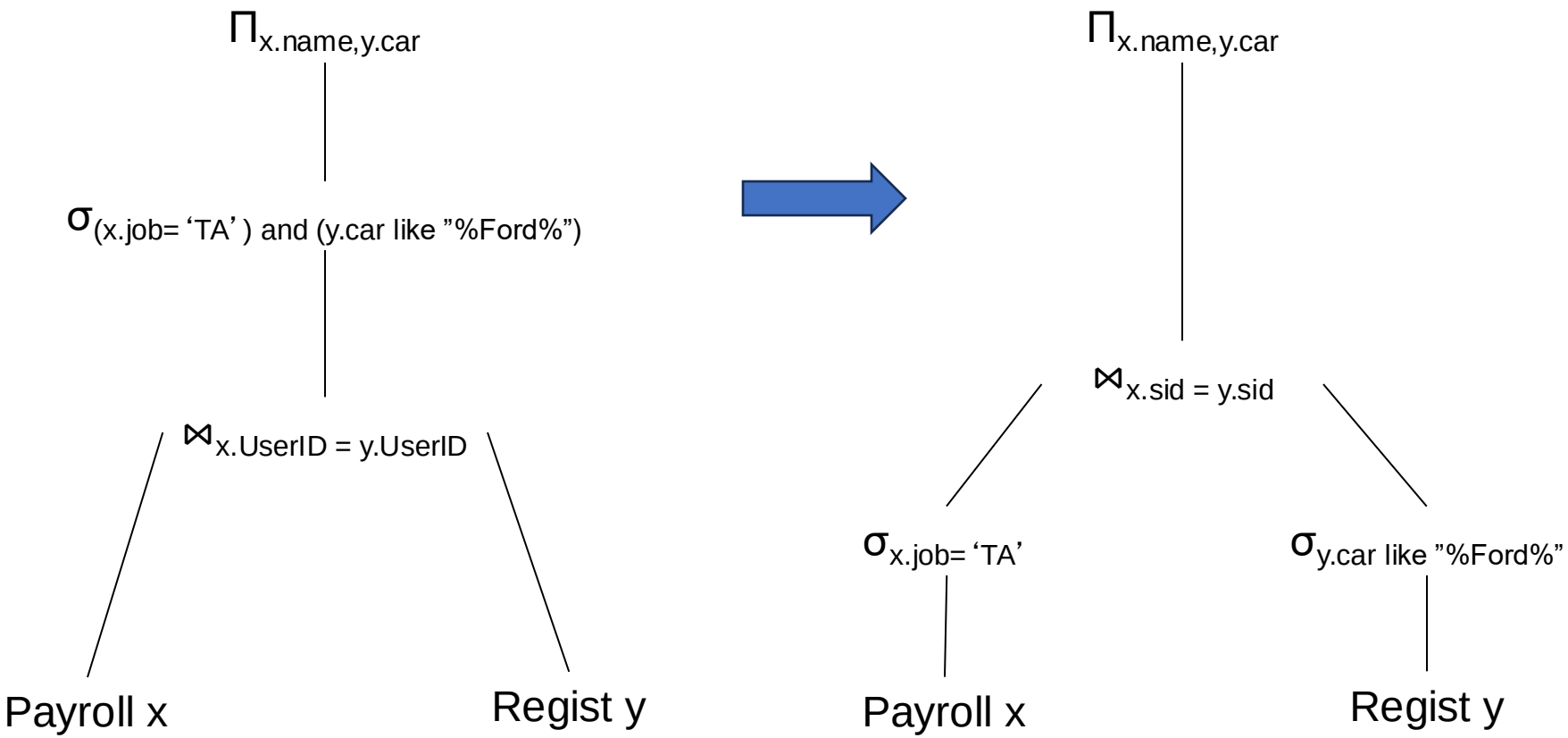
$\sigma_C(R \bowtie S) = \sigma_C(R) \bowtie S$ when C refers only to R



Rewrite Rules: Predicate Pushdown

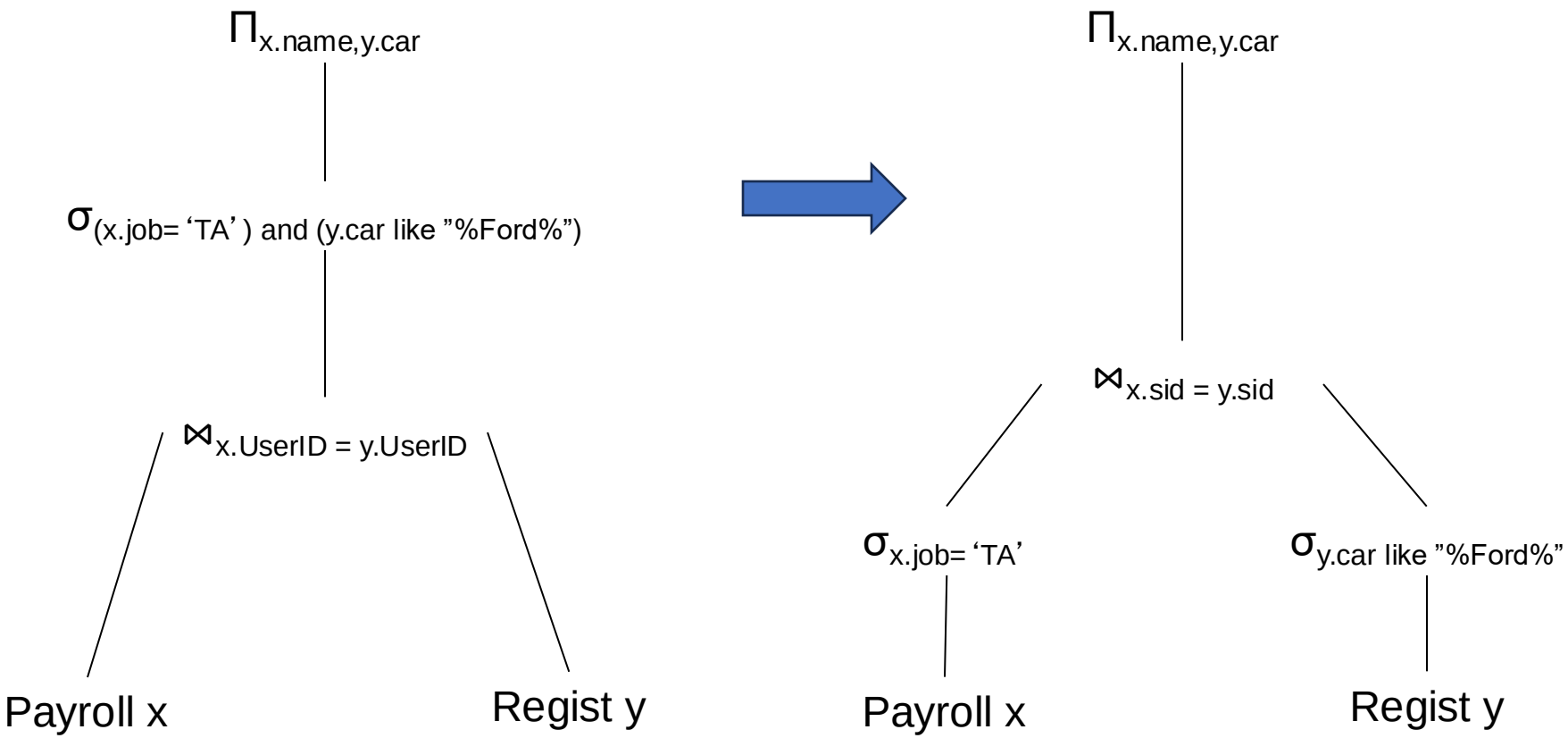


Rewrite Rules: Predicate Pushdown

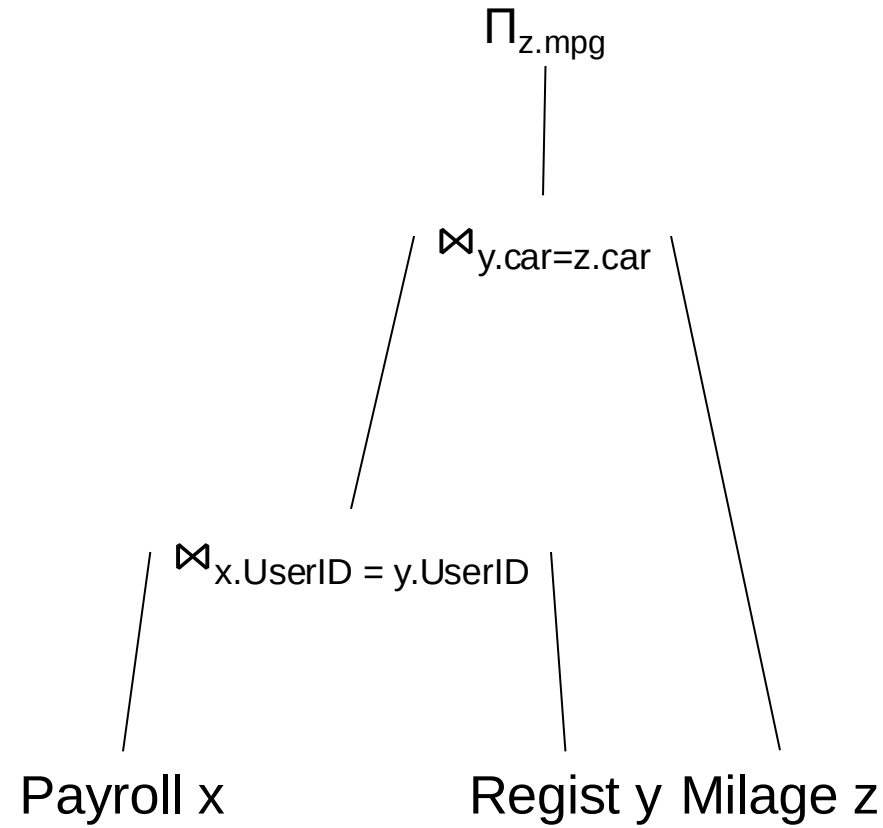


Rewrite Rules: Predicate Pushdown

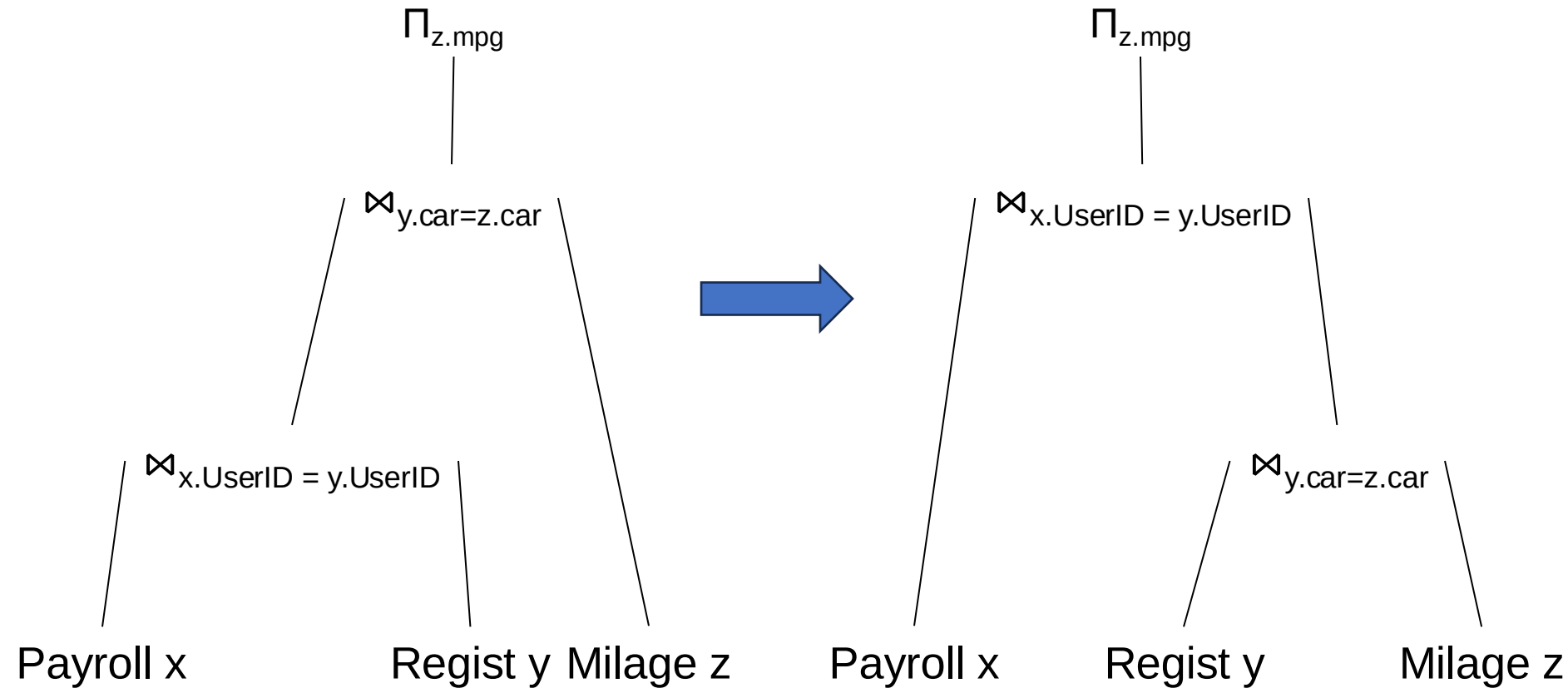
$$\sigma_{C1 \text{ and } C2}(R \bowtie S) = \sigma_{C1}(\sigma_{C2}(R \bowtie S))$$



Rewrite Rules: Join Associativity

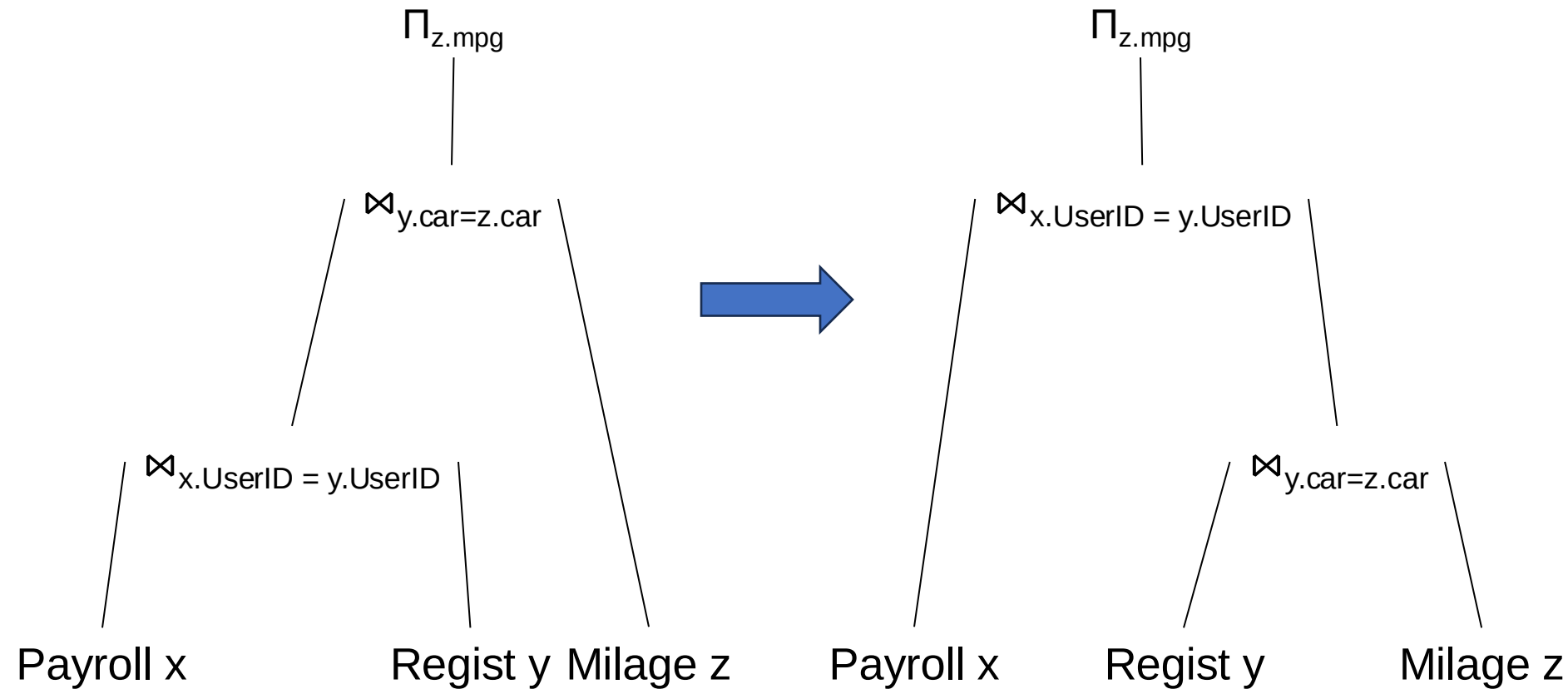


Rewrite Rules: Join Associativity



Rewrite Rules: Join Associativity

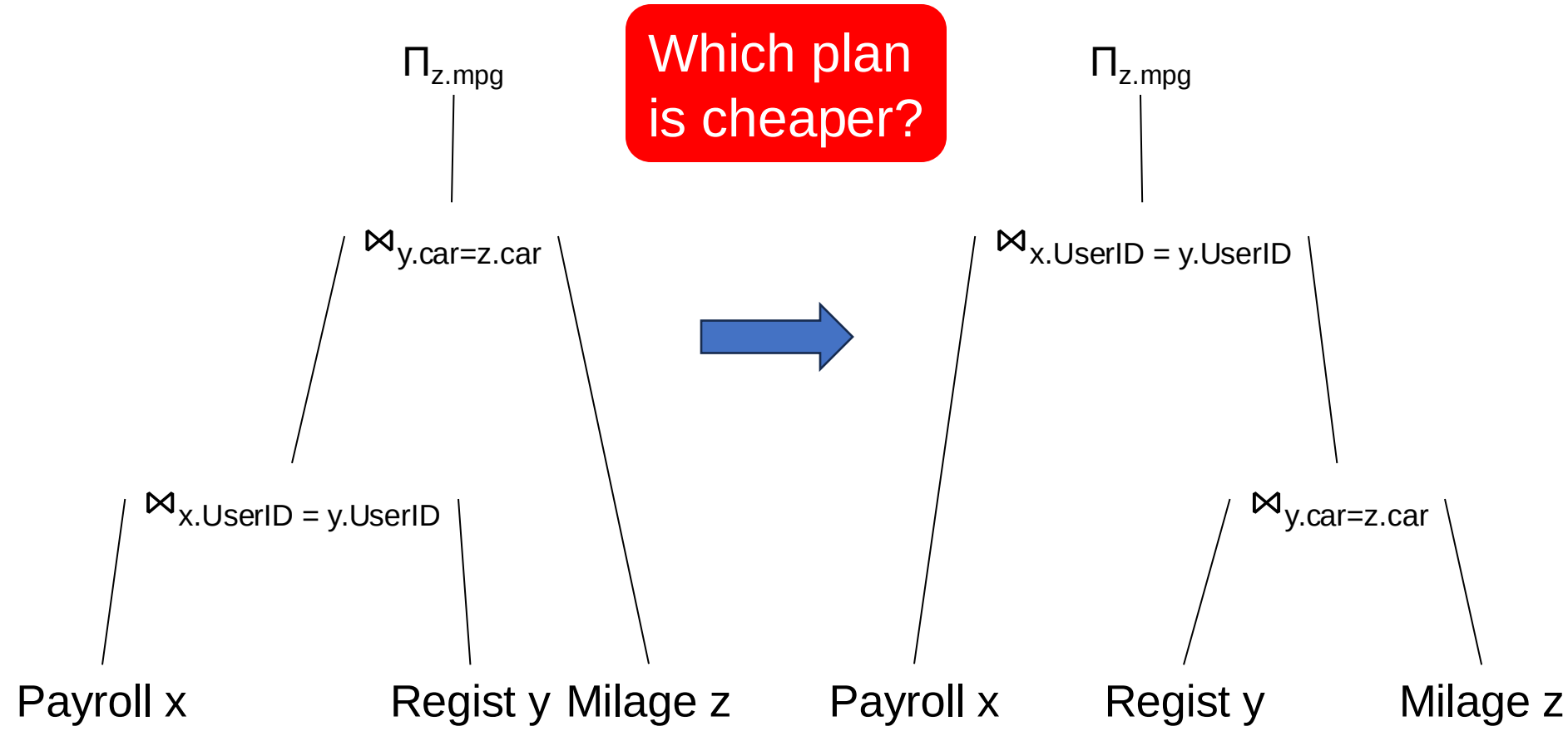
$$(R \bowtie S) \bowtie T = R \bowtie (S \bowtie T)$$



Rewrite Rules: Join Associativity

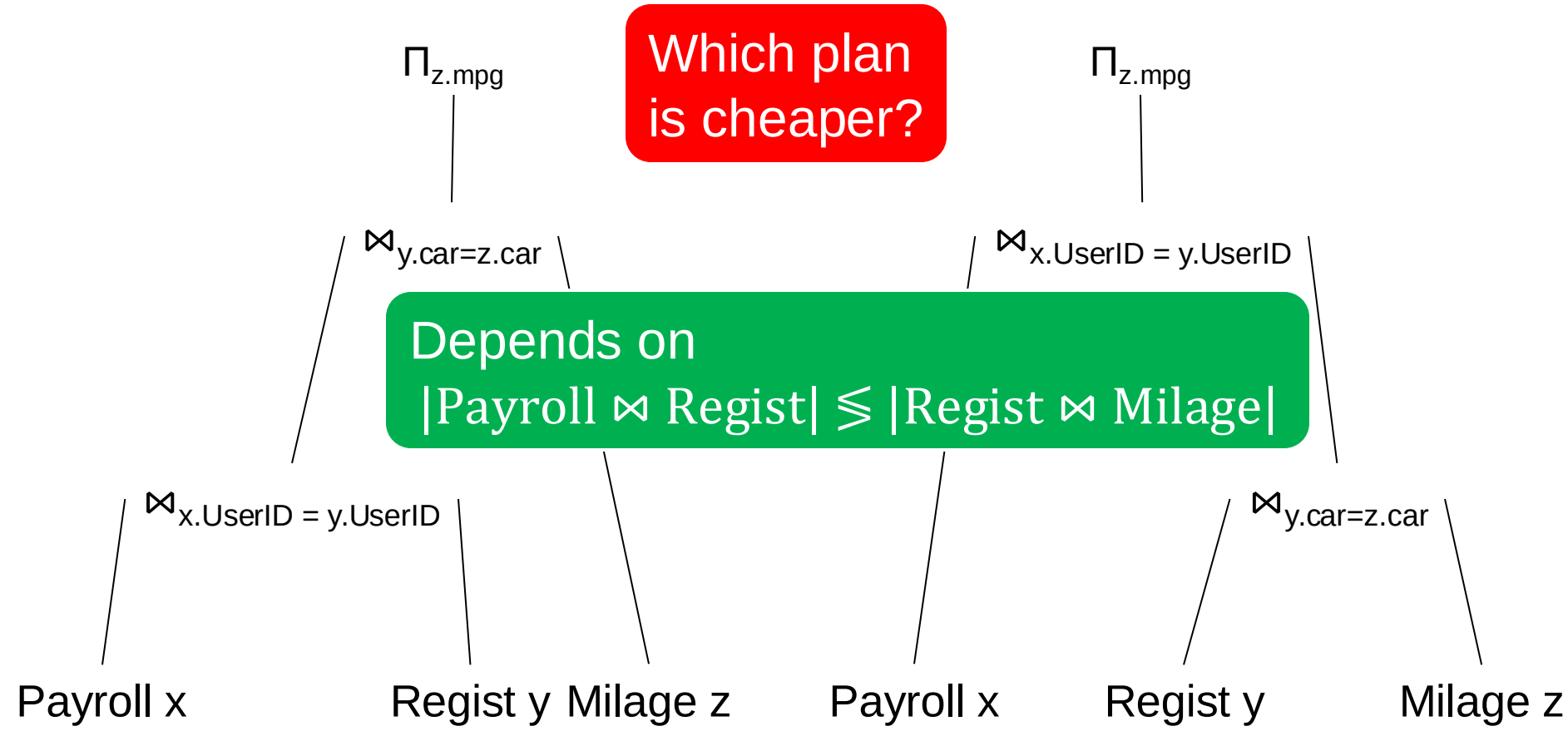
$$(R \bowtie S) \bowtie T = R \bowtie (S \bowtie T)$$

Which plan is cheaper?

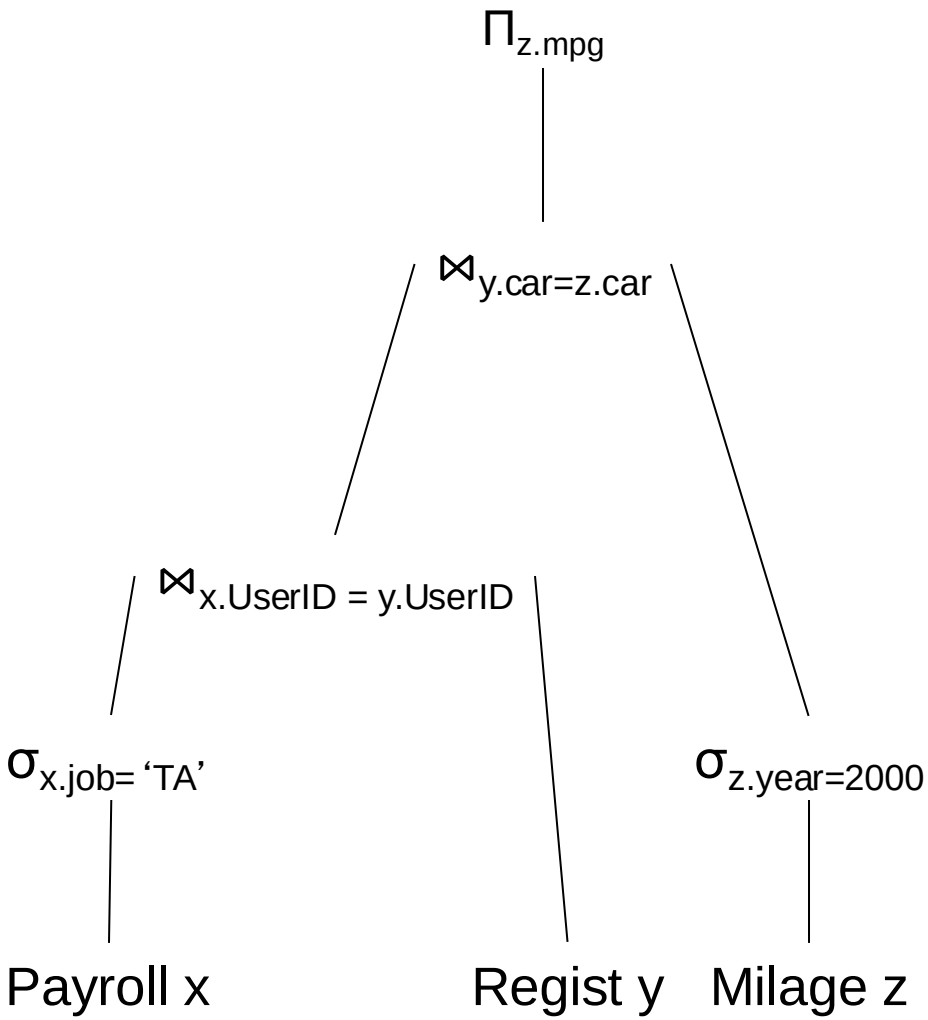


Rewrite Rules: Join Associativity

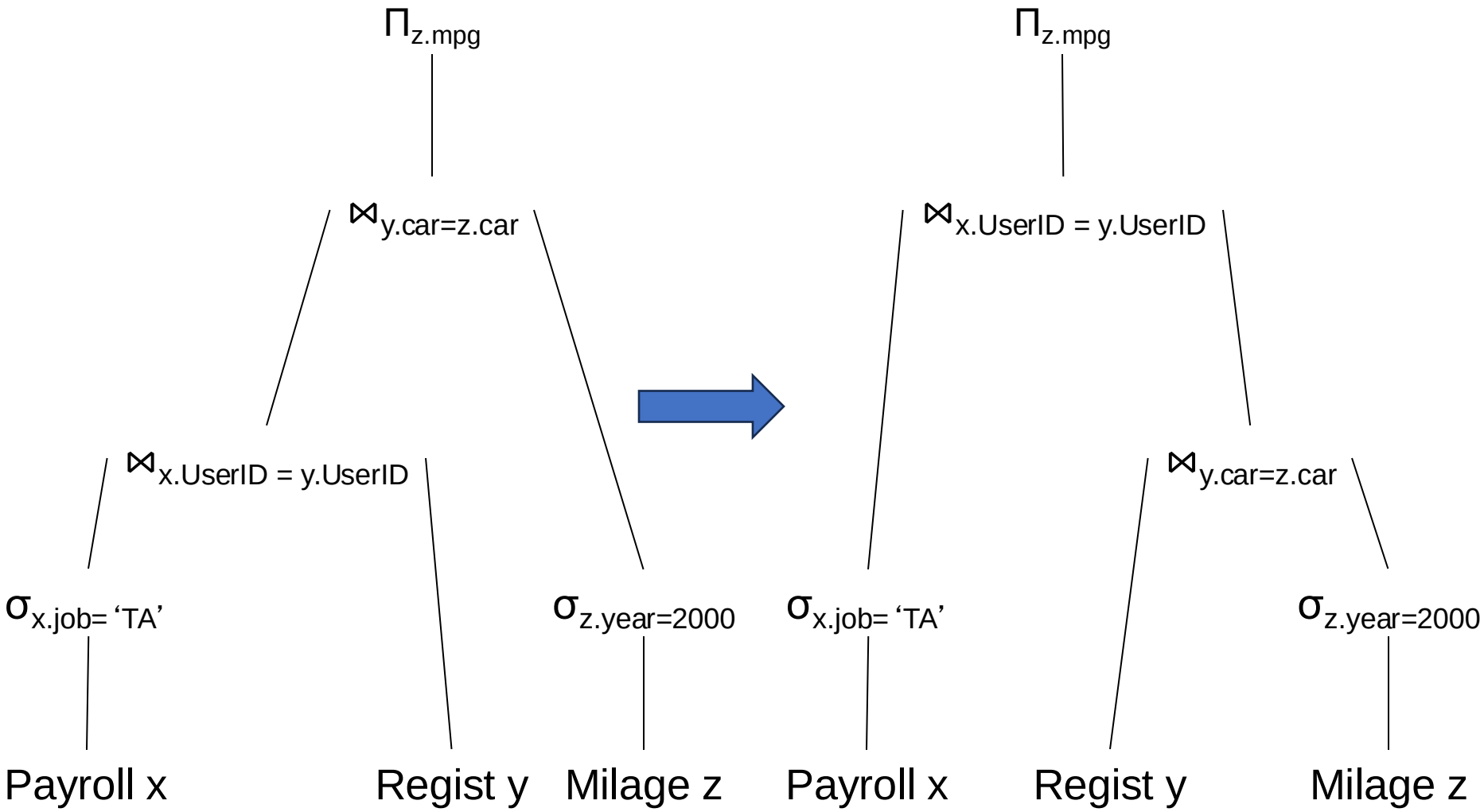
$$(R \bowtie S) \bowtie T = R \bowtie (S \bowtie T)$$



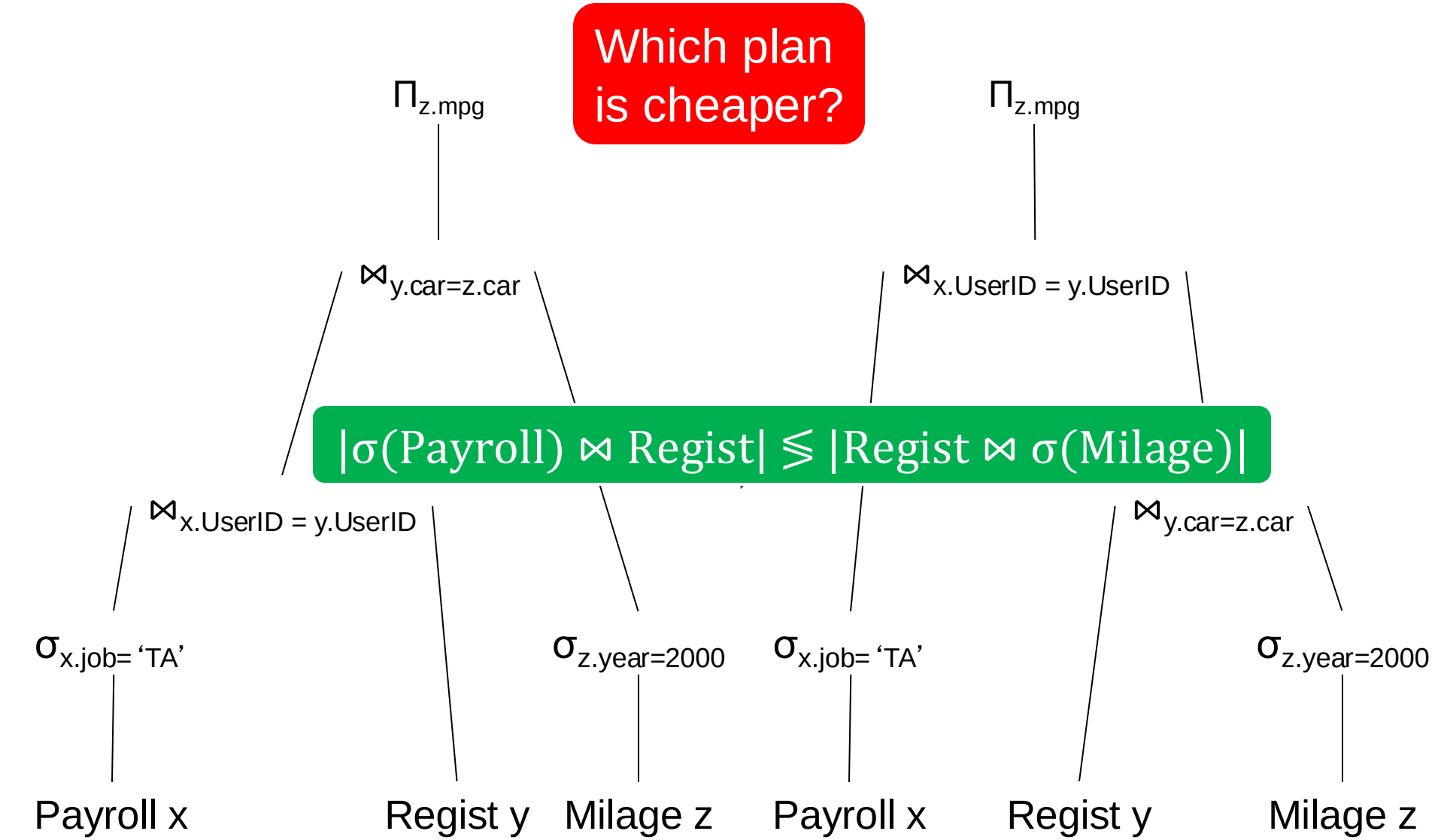
Rewrite Rules: Join Associativity



Rewrite Rules: Join Associativity



Rewrite Rules: Join Associativity



Practice

- Optimizers have a fixed list of rewrite rules
- E.g. SQL Server: about 500 rules
- Rules become complex and specialized

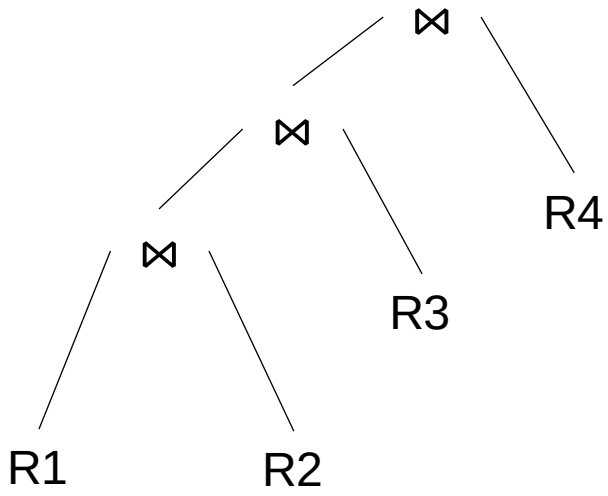
Search Space Challenges

Search space is huge!

Compromises:

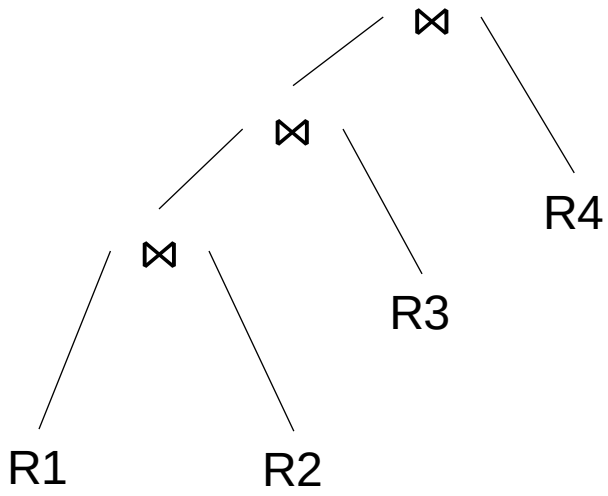
- Left-deep plans or right-deep plans
- Plans without cartesian products

Left-Deep-, Right-Deep-, Bushy-Plans

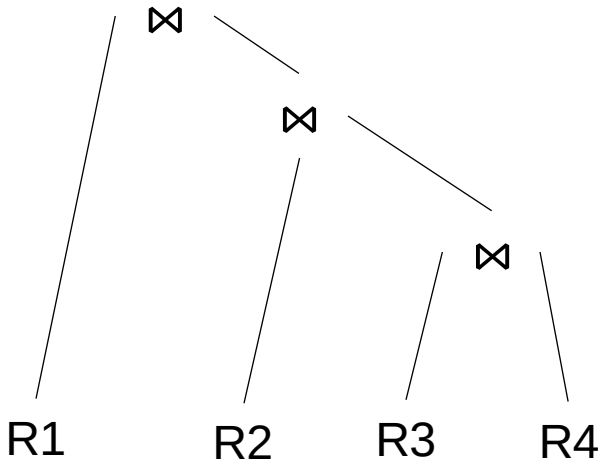


Left Deep

Left-Deep-, Right-Deep-, Bushy-Plans

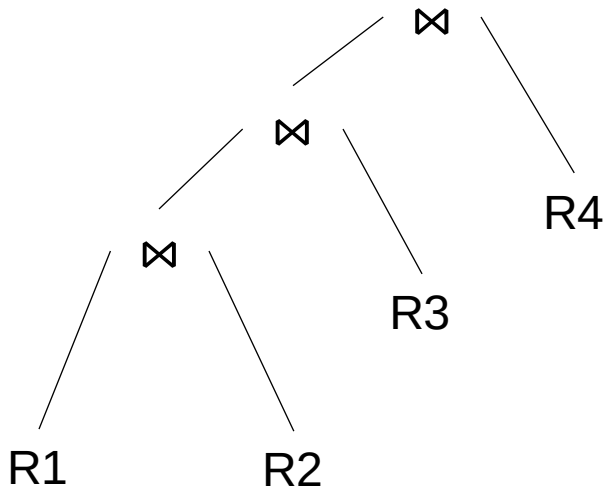


Left Deep

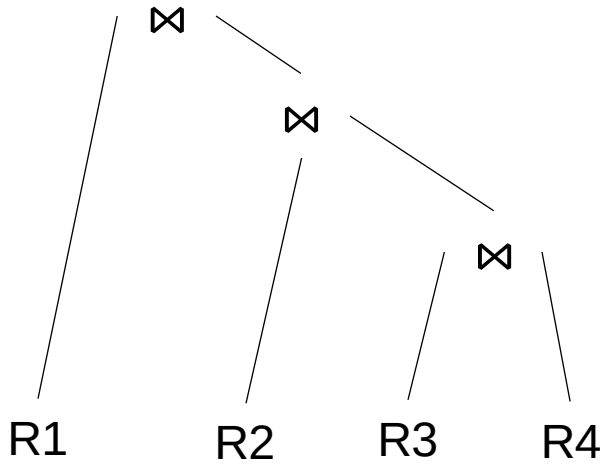


Right Deep

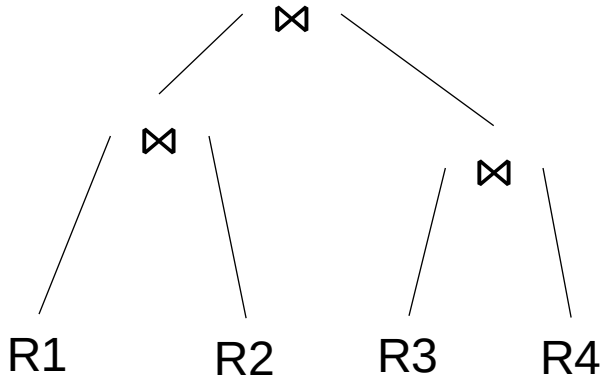
Left-Deep-, Right-Deep-, Bushy-Plans



Left Deep



Right Deep



Bushy

Left-Deep-, Right-Deep-, Bushy-Plans

Join n relations R_1, R_2, …, R_n

- How many left-deep plans are there?
- How many right-deep plans?
- How many bushy plans?

Left-Deep-, Right-Deep-, Bushy-Plans

Join n relations R_1, R_2, …, R_n

- How many left-deep plans are there? n!
- How many right-deep plans? n!
- How many bushy plans?

Left-Deep-, Right-Deep-, Bushy-Plans

Join n relations R_1, R_2, …, R_n

- How many left-deep plans are there? n!
- How many right-deep plans? n!
- How many bushy plans? n! × C_n

  where $C_{n-1} = \frac{1}{n}\binom{2n}{n-1}$ is the Catalan number

Left-Deep-, Right-Deep-, Bushy-Plans

Join n relations R_1, R_2, …, R_n

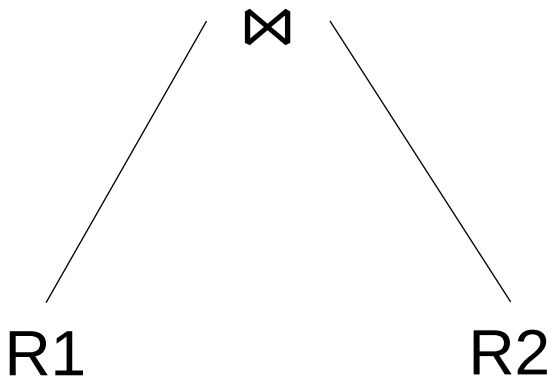
- How many left-deep plans are there? n!
- How many right-deep plans? n!
- How many bushy plans? n! × C_n

  where $C_{n-1} = \frac{1}{n}\binom{2n}{n-1}$ is the Catalan number

Many DBMS use only left- or right-deep plans

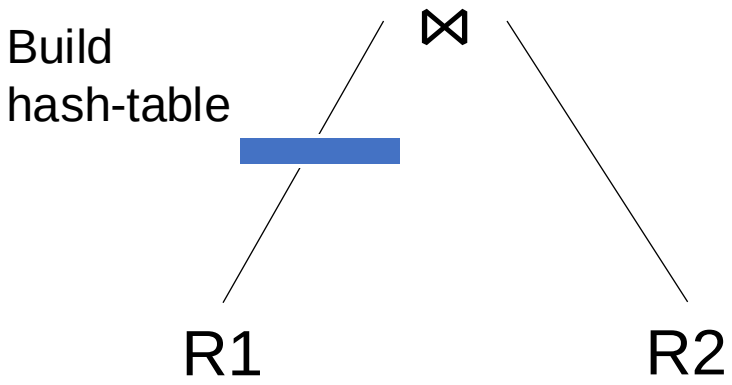
Pipelining v.s. Blocking

Hash-Join



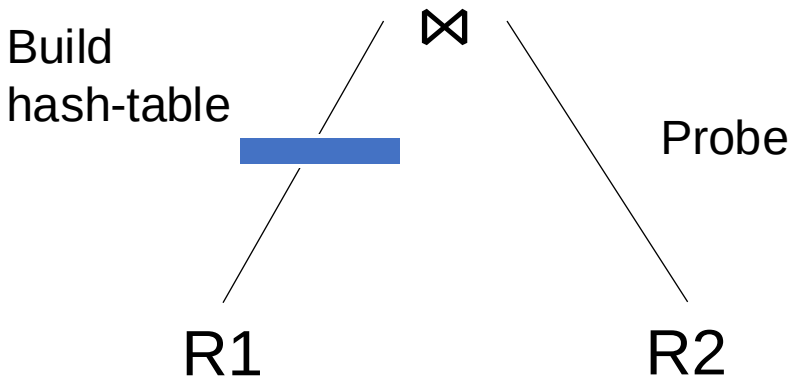
Pipelining v.s. Blocking

Hash-Join



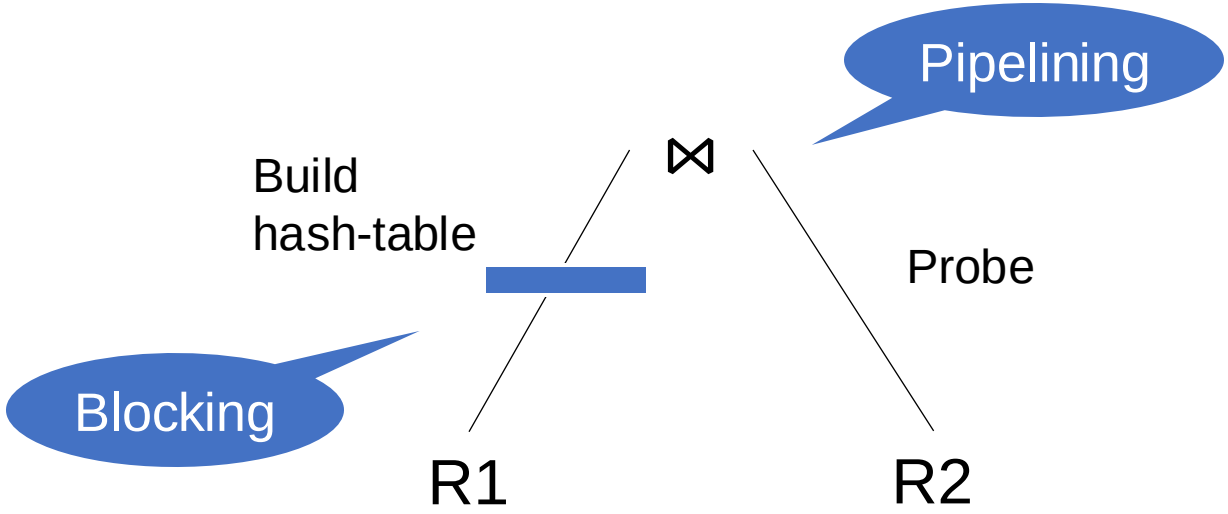
Pipelining v.s. Blocking

Hash-Join



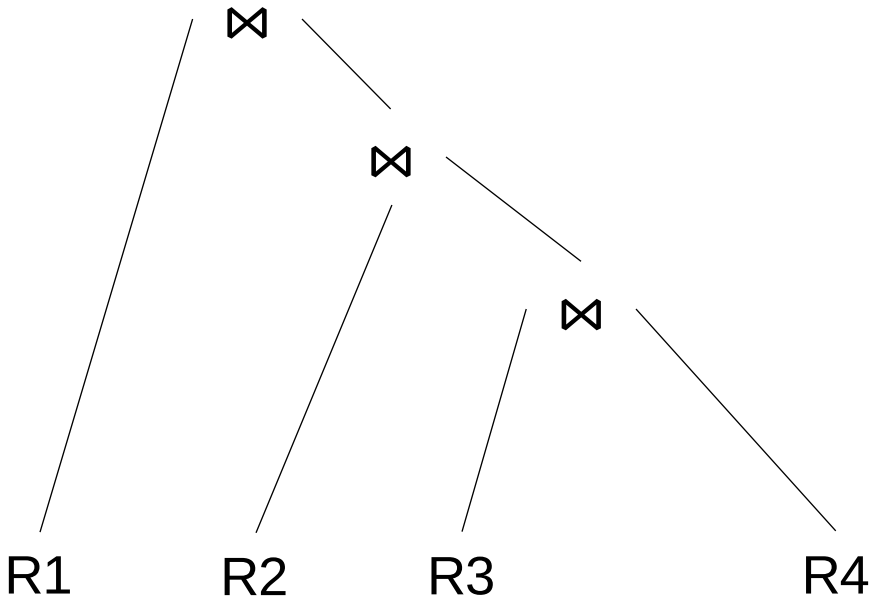
Pipelining v.s. Blocking

Hash-Join



Pipelining v.s. Blocking

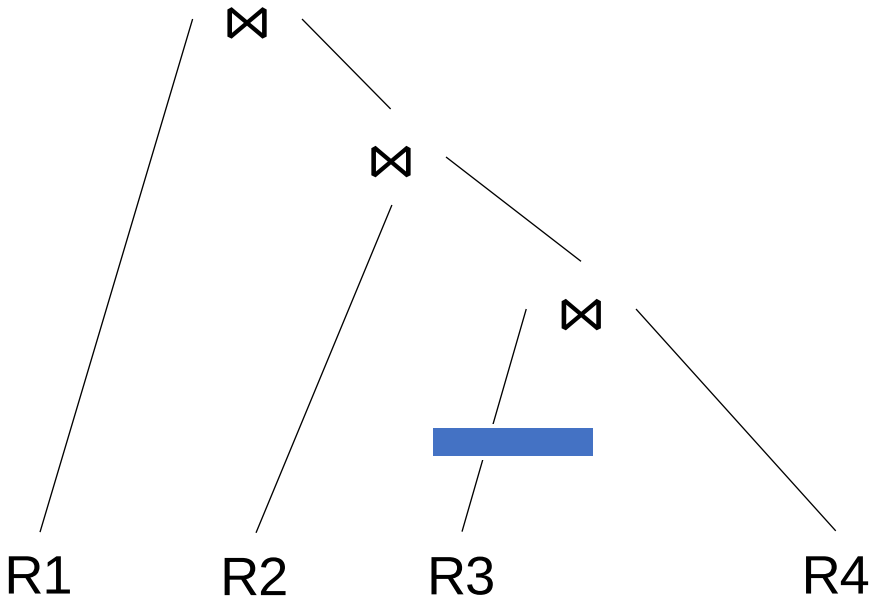
Hash-Join



Pipelining v.s. Blocking

Hash-Join

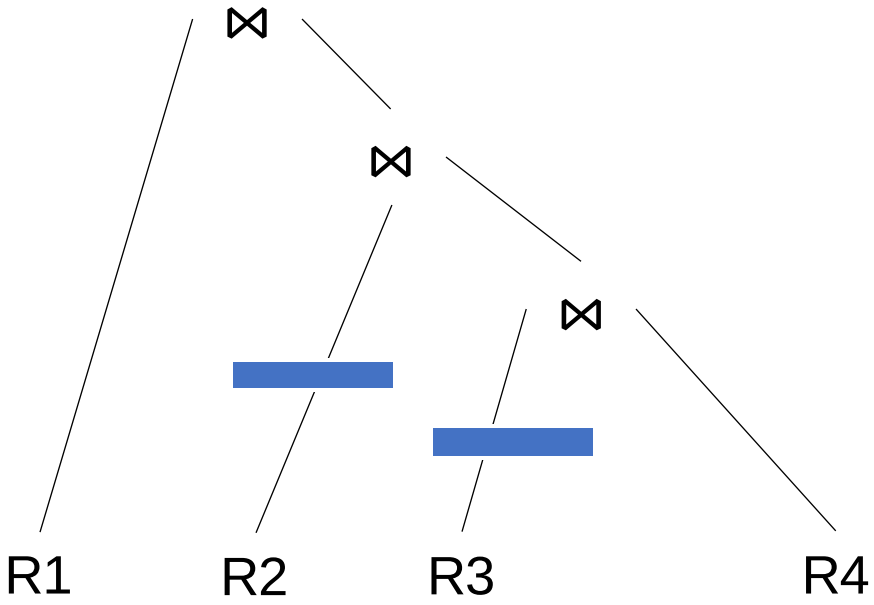
Build
hash-tables



Pipelining v.s. Blocking

Hash-Join

Build
hash-tables

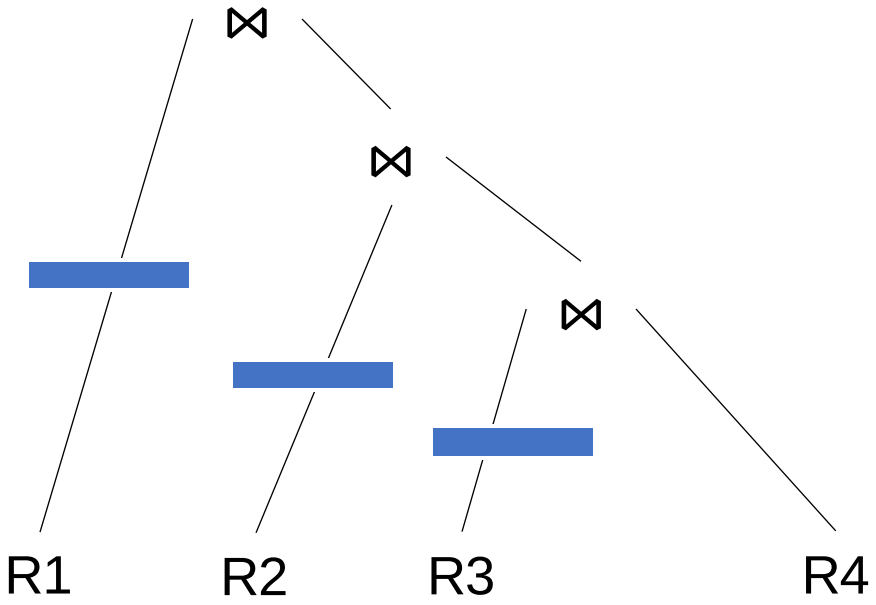


Pipelining v.s. Blocking

Hash-Join

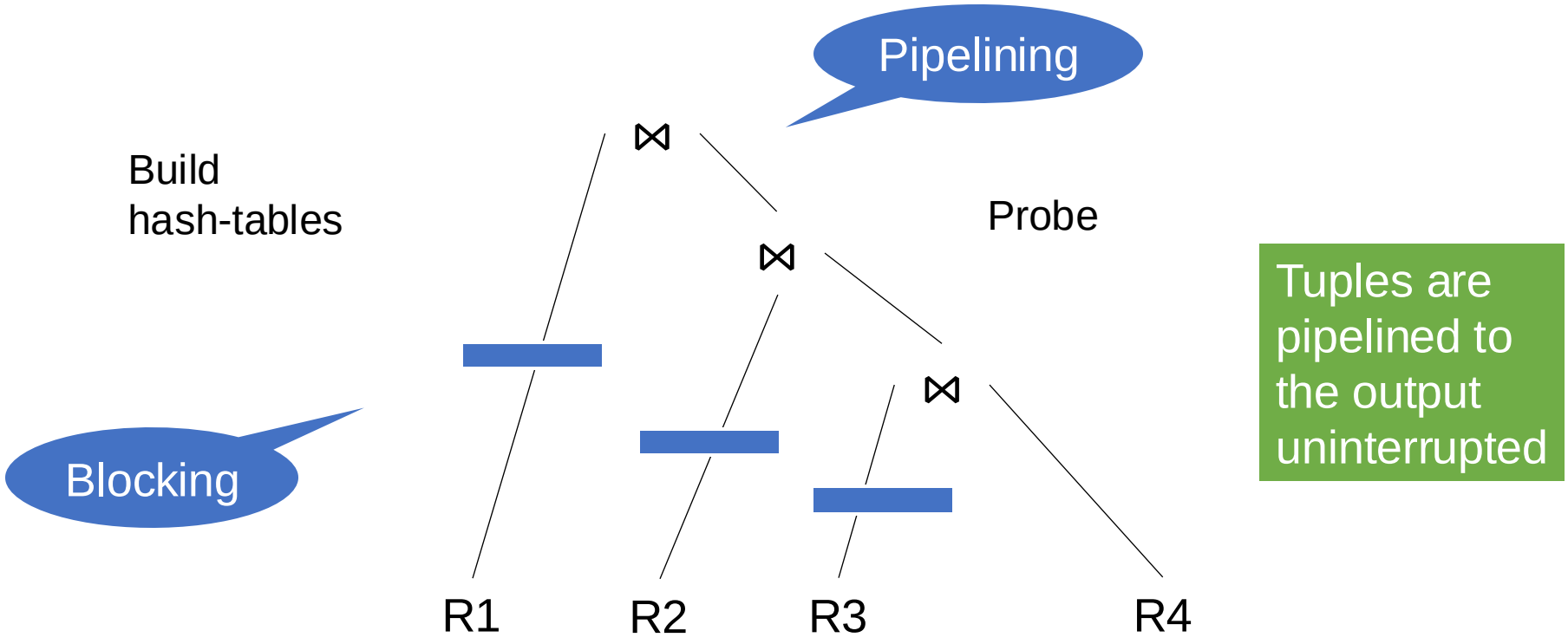
Build
hash-tables

Blocking



Pipelining v.s. Blocking

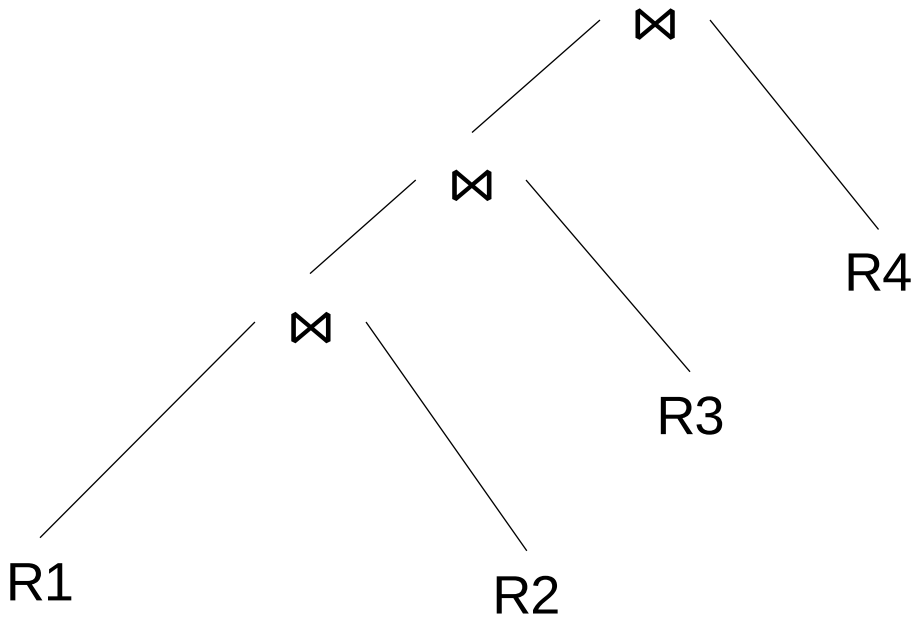
Hash-Join



Tuples are pipelined to the output uninterrupted

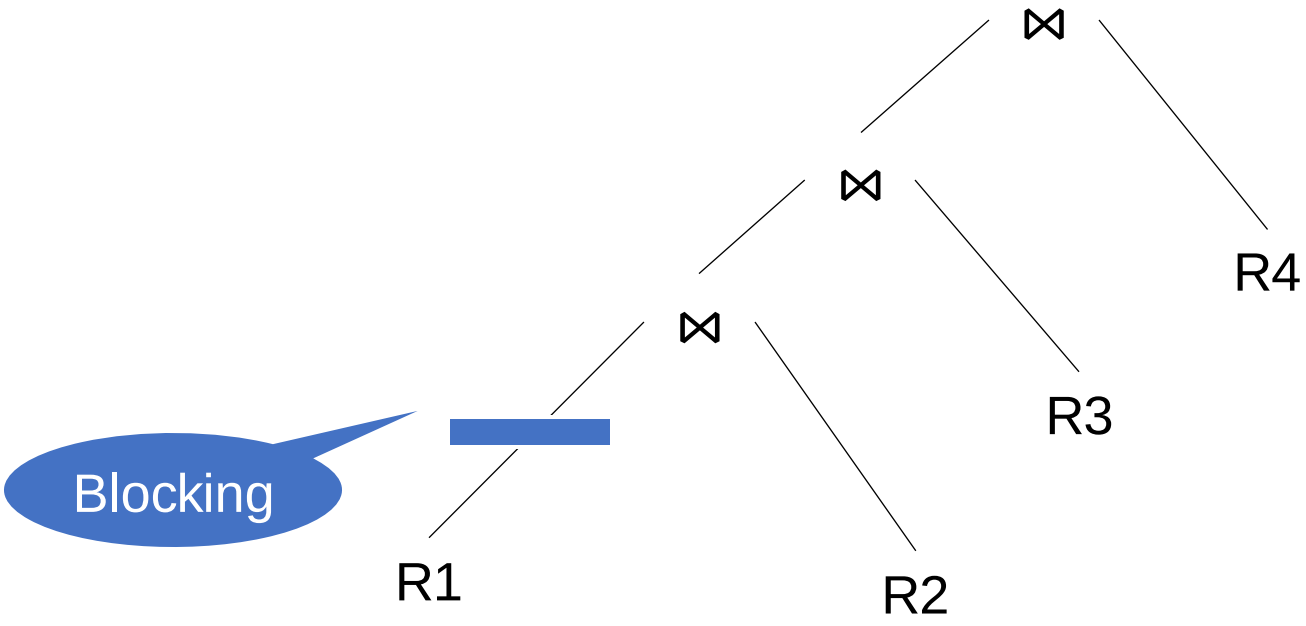
Pipelining v.s. Blocking

Hash-Join



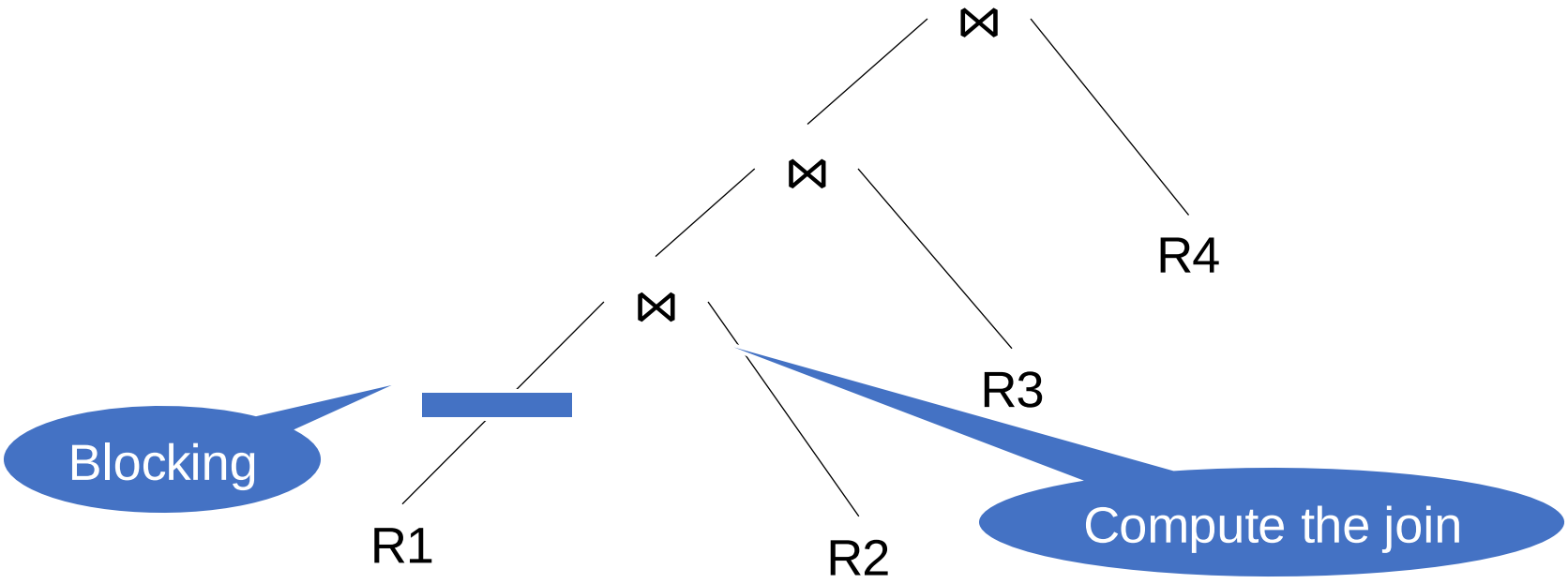
Pipelining v.s. Blocking

Hash-Join



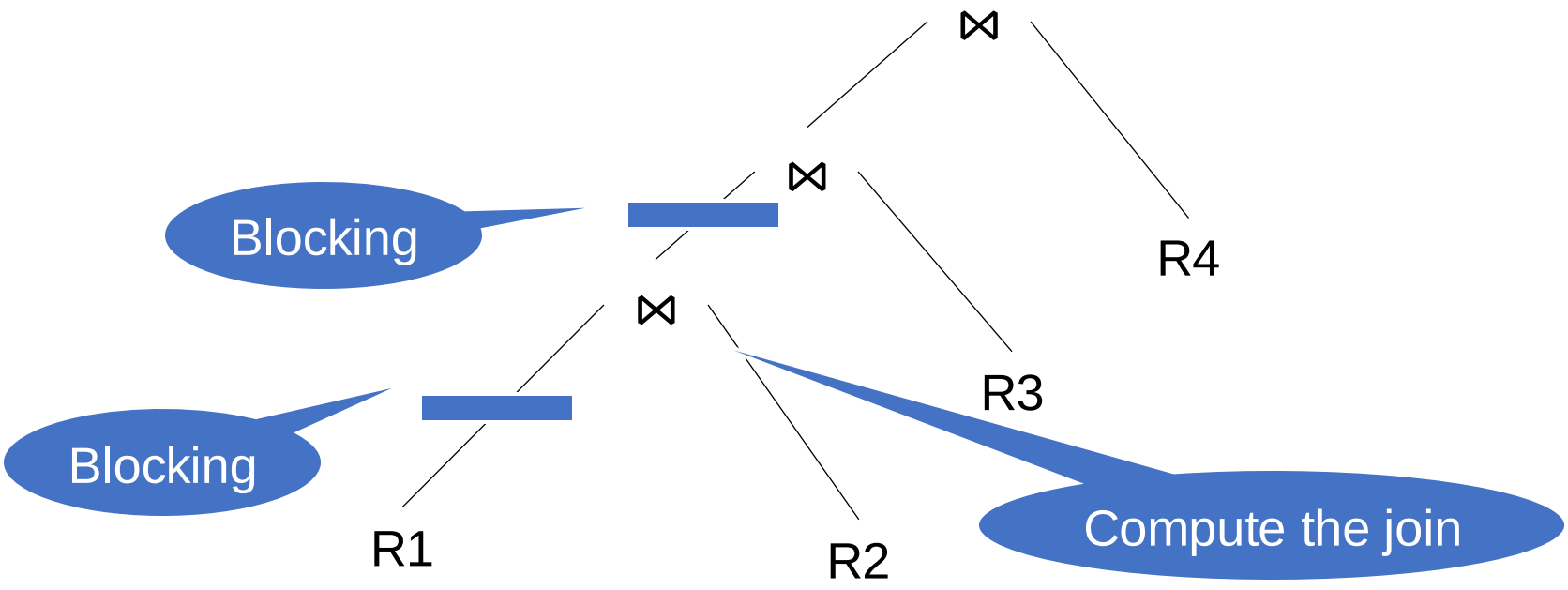
Pipelining v.s. Blocking

Hash-Join



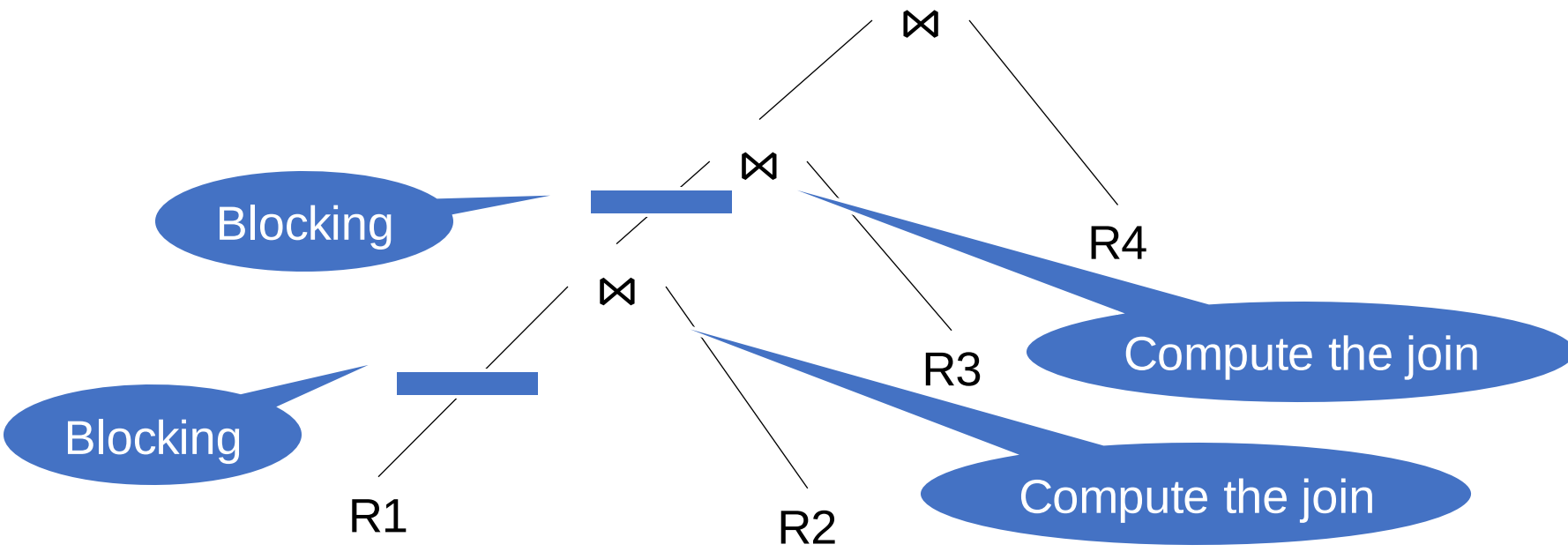
Pipelining v.s. Blocking

Hash-Join



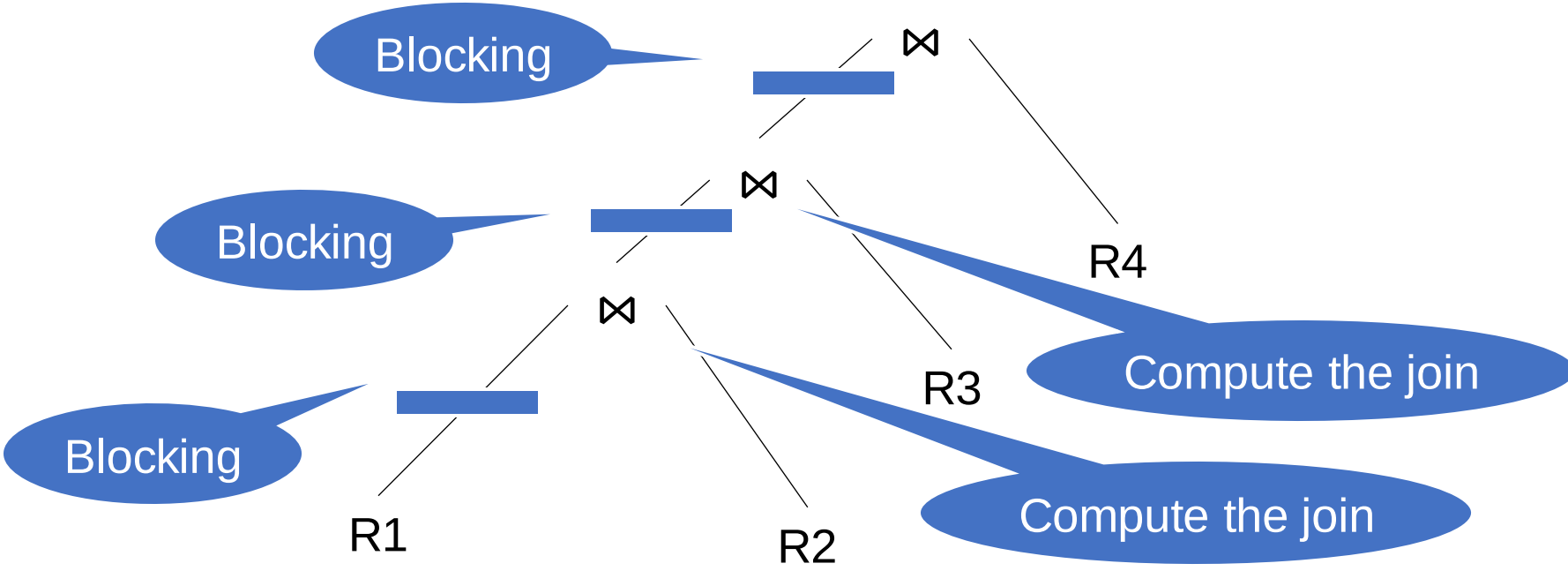
Pipelining v.s. Blocking

Hash-Join



Pipelining v.s. Blocking

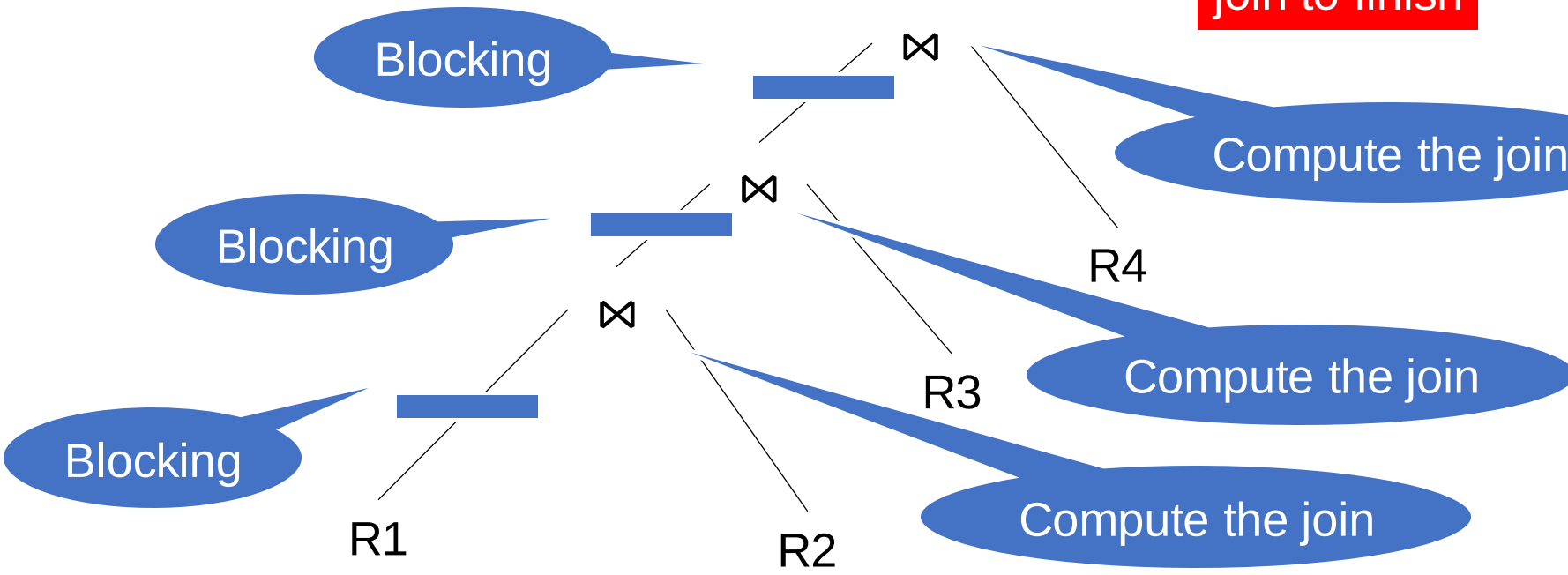
Hash-Join



Pipelining v.s. Blocking

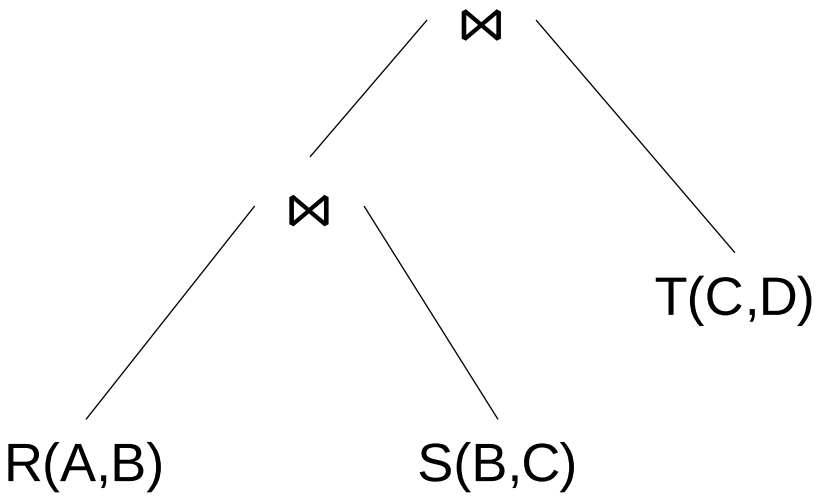
Hash-Join

Each build phase must wait for the previous join to finish



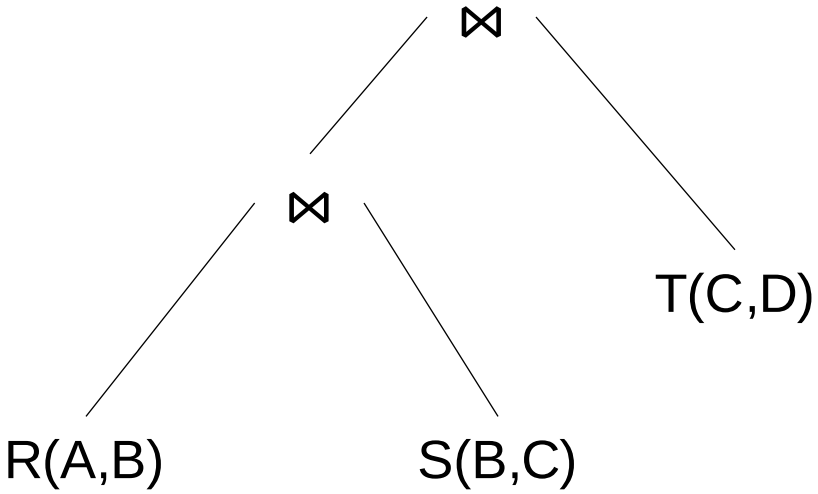
With or Without Cartesian Products

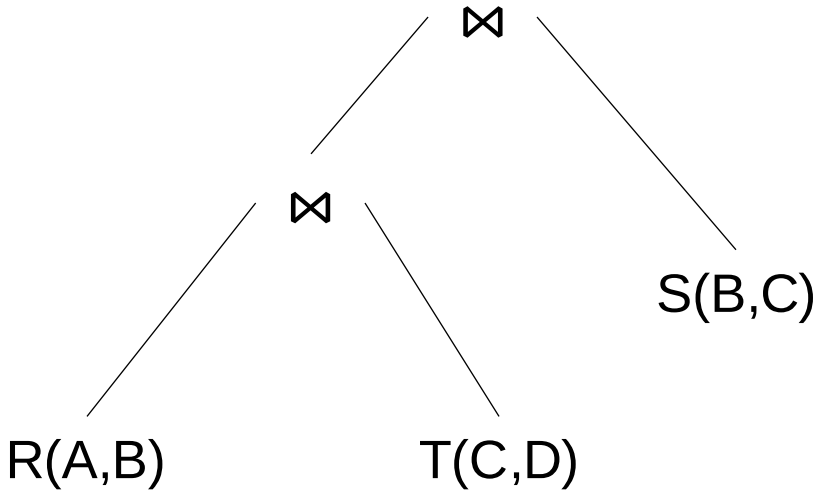
All joins are natural joins:



With or Without Cartesian Products

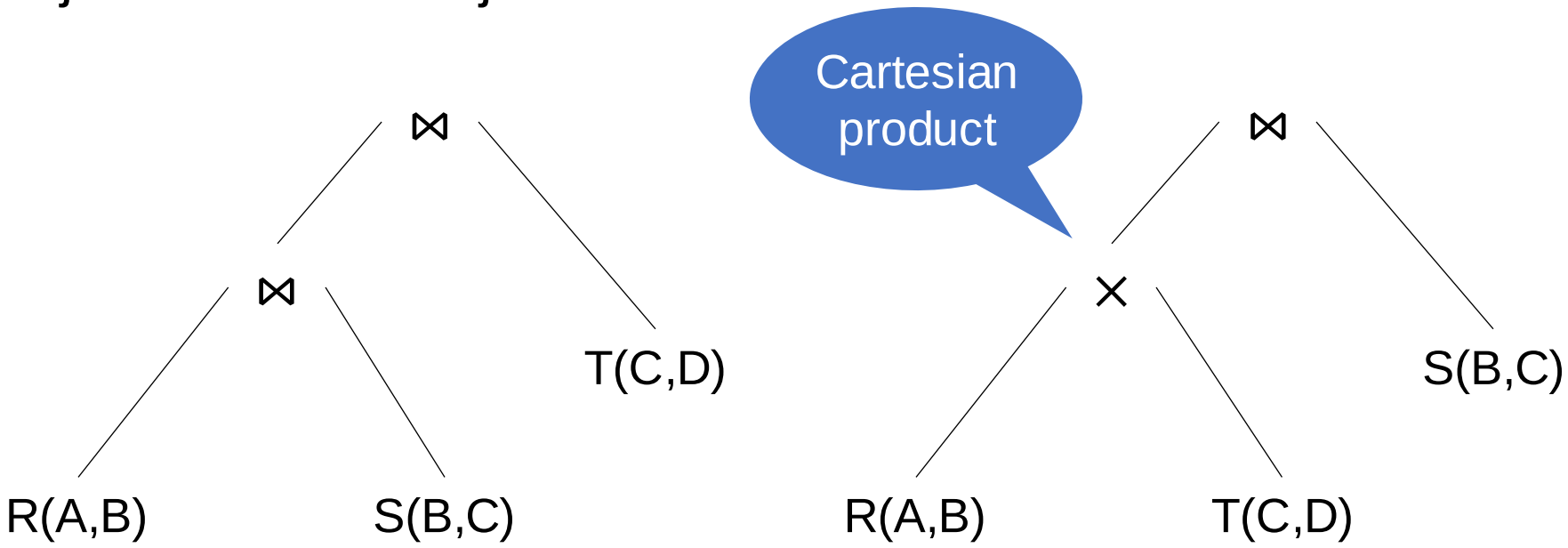
All joins are natural joins:





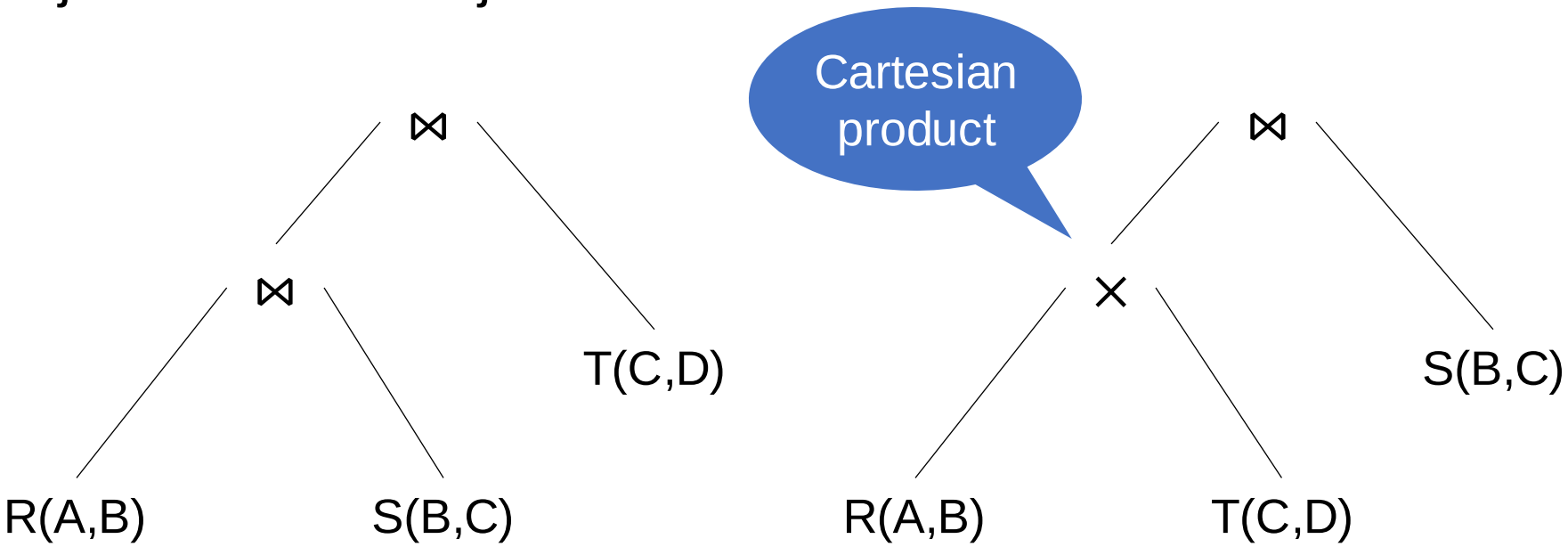
With or Without Cartesian Products

All joins are natural joins:



With or Without Cartesian Products

All joins are natural joins:

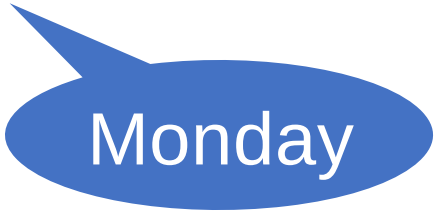


Most optimizers avoid cartesian products

Discussion

- The optimizer applies rewrite rules to generate alternative query plans

- In order to choose the cheapest plan, it needs to estimate the cardinality of various plans



Access Path Selection

Access Path

Access path: implements a selection $\sigma_P(R)$:

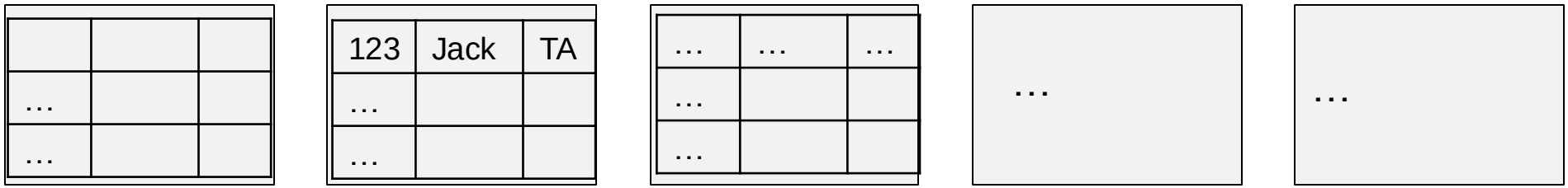
- A file scan, or
- An index-based selection

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

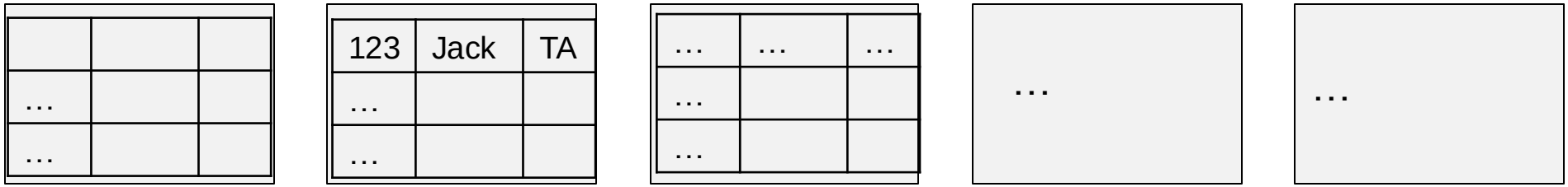


Payroll data file

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

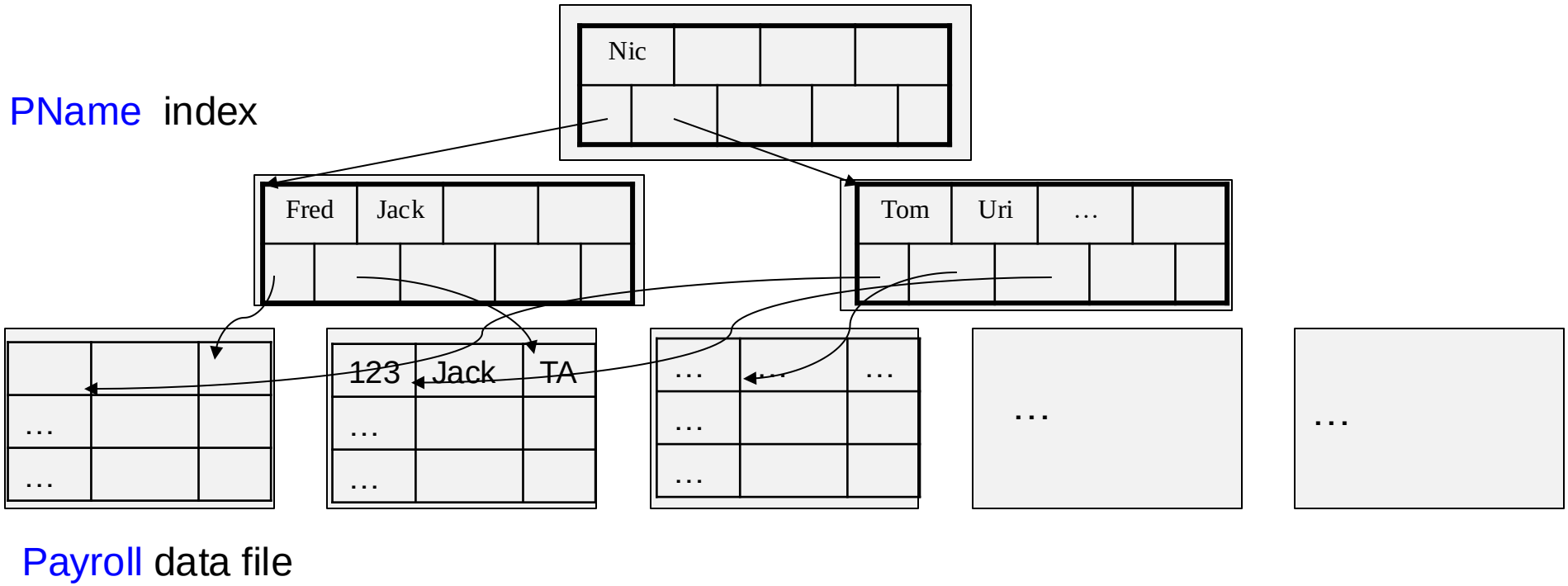


Payroll data file

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```



Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

```
SELECT *
FROM Payroll
WHERE name = 'Allison';
```

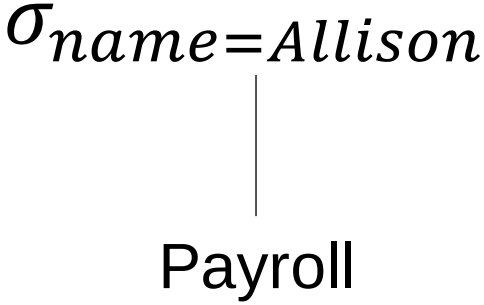
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

```
SELECT *
FROM Payroll
WHERE name = 'Allison';
```



Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

```
SELECT *
FROM Payroll
WHERE name = 'Allison';
```

$\sigma_{name=Allison}$

Payroll

Option 1: Index-based selection

Cost = ???

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

```
SELECT *
FROM Payroll
WHERE name = 'Allison';
```

$\sigma_{name=Allison}$

Payroll

Option 1: Index-based selection

Cost = T / V = 50

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

```
SELECT *
FROM Payroll
WHERE name = 'Allison';
```

$\sigma_{name=Allison}$

Payroll

Option 1: Index-based selection

Cost = T / V = 50

Option 2: Table scan

Cost = ???

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

```
SELECT *
FROM Payroll
WHERE name = 'Allison';
```

$\sigma_{name=Allison}$

Payroll

Option 1: Index-based selection

Cost = T / V = 50

Option 2: Table scan

Cost = B = 5000

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

```
SELECT *
FROM Payroll
WHERE name = 'Allison';
```

$\sigma_{name=Allison}$

Payroll

Option 1: Index-based selection

Cost = T / V = 50

Best access path

Option 2: Table scan

Cost = B = 5000

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000

```
SELECT *
FROM Payroll
WHERE name = 'Allison';
```

```
SELECT *
FROM Payroll x, Regist y
WHERE x.UserID = y.UserID;
```

Indexes are not magic!

Access Path Selection

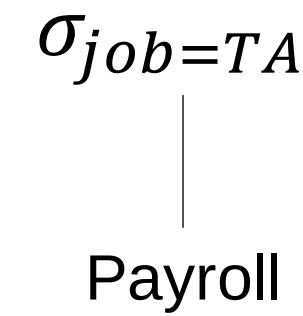
```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
CREATE INDEX Pjob on Payroll(job)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000
V(Payroll, job) = 25

```
SELECT *
FROM Payroll
WHERE job = 'TA';
```

Option 1: Index-based selection



Option 2: Table scan

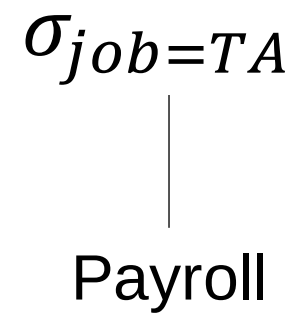
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
CREATE INDEX Pjob on Payroll(job)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000
V(Payroll, job) = 25

```
SELECT *
FROM Payroll
WHERE job = 'TA';
```



Option 1: Index-based selection

Cost = T / V = 40000

Best access path

Option 2: Table scan

Cost = B = 5000

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
CREATE CLUSTERED INDEX Pjob on Payroll(job)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000
V(Payroll, job) = 25

```
SELECT *
FROM Payroll
WHERE job = 'TA';
```

$\sigma_{job=TA}$

|

Payroll

Option 1: Index-based selection

Cost = ???

Option 2: Table scan

Cost = B = 5000

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name)
CREATE CLUSTERED INDEX Pjob on Payroll(job)
```

T(Payroll) = 1000000
B(Payroll) = 5000
V(Payroll, name) = 20000
V(Payroll, job) = 25

```
SELECT *
FROM Payroll
WHERE job = 'TA';
```

$\sigma_{job=TA}$

|

Payroll

Option 1: Index-based selection

Cost = B / V = 200

Best access path

Option 2: Table scan

Cost = B = 5000

Multiple Indexes

- DB administrator can create multiple indexes
- But only one can be clustered…
- … and a selection can only use only one index

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

Payroll:

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

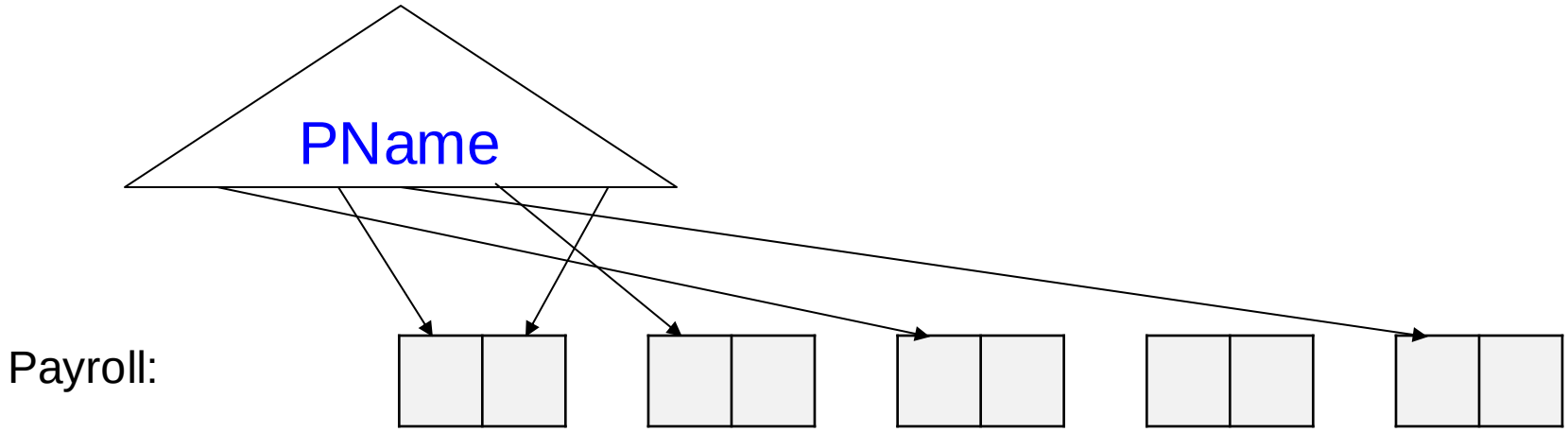
```
CREATE INDEX Pname on Payroll(name);
```

Payroll:

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

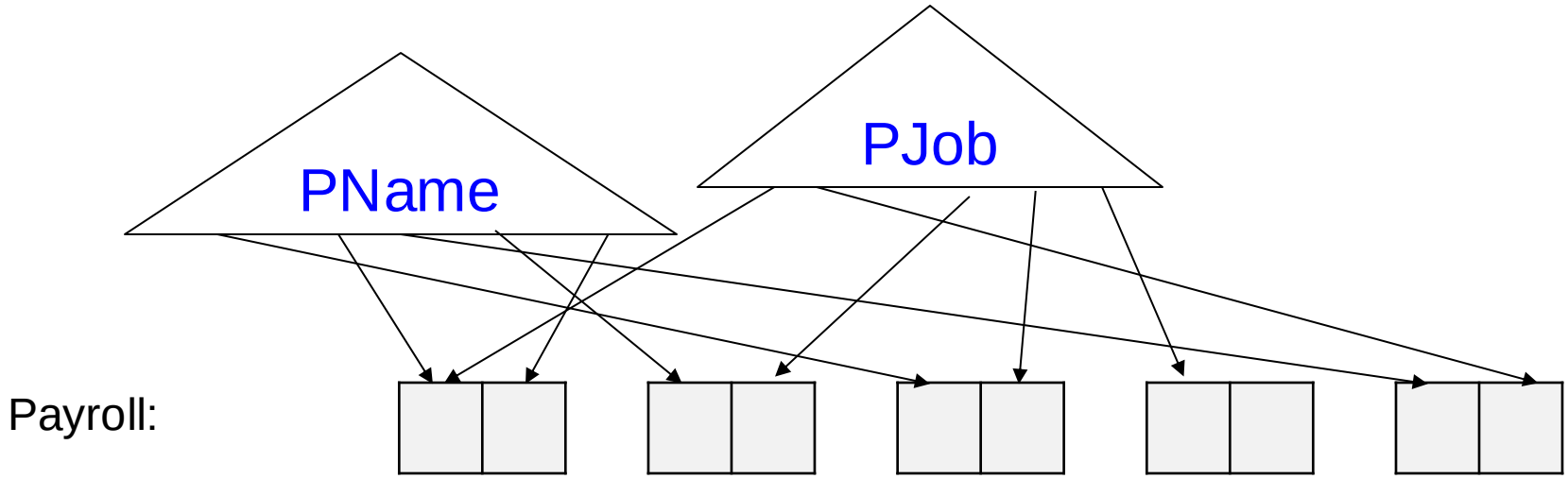
```
CREATE INDEX Pname on Payroll(name);
```



Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

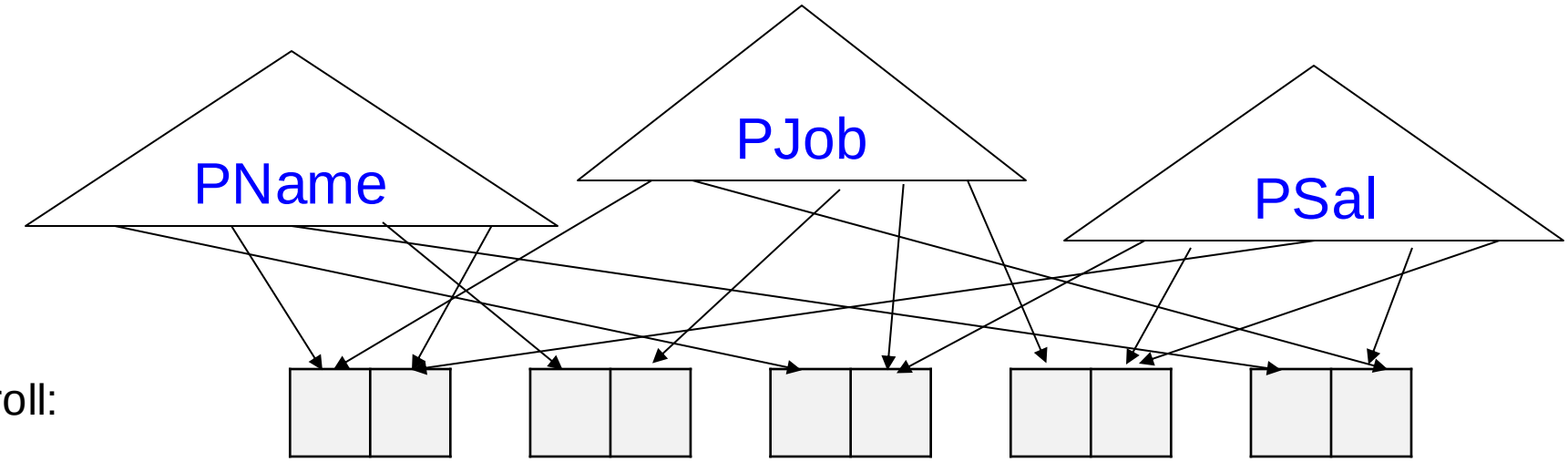
```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
```



Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

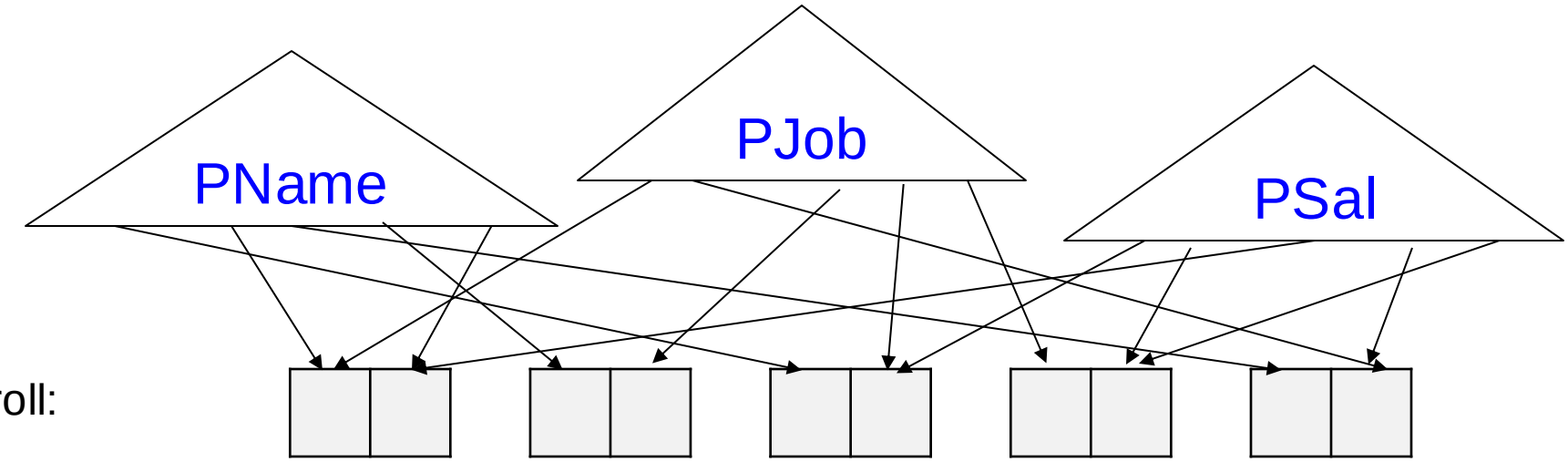


Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```



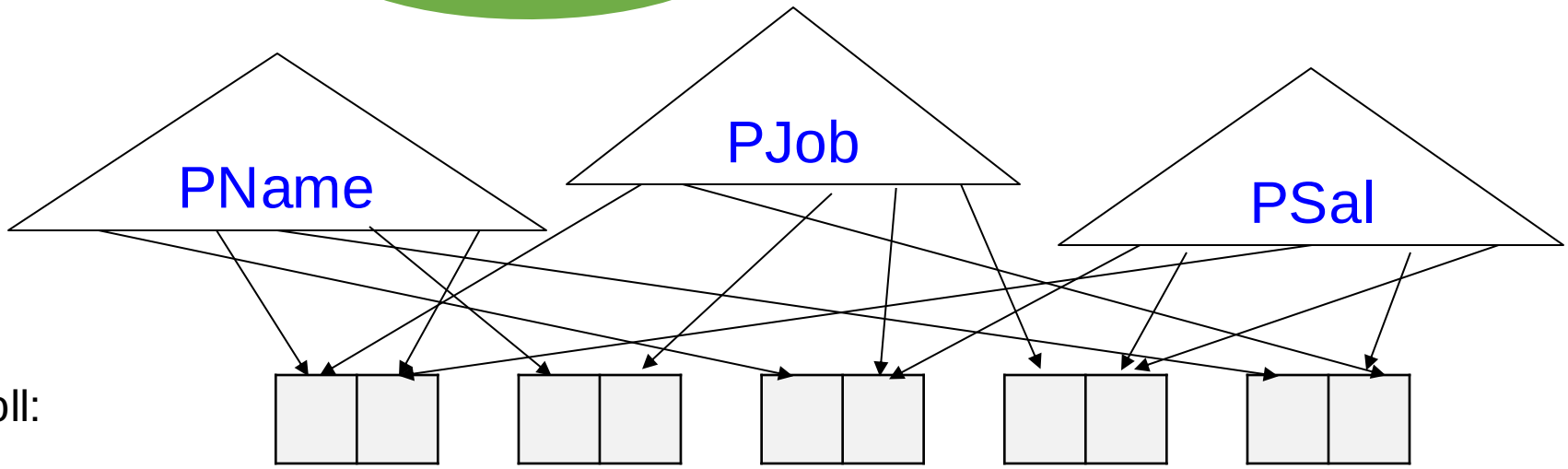
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Access path:
use Pjob index



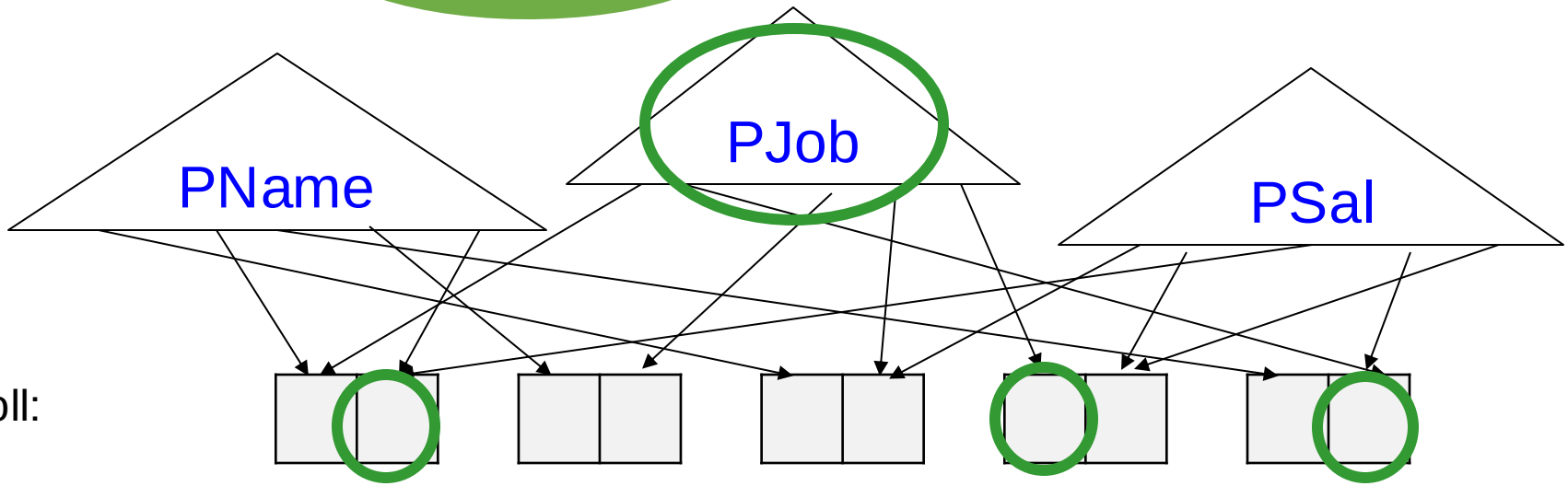
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Access path:
use Pjob index



Access Path Selection

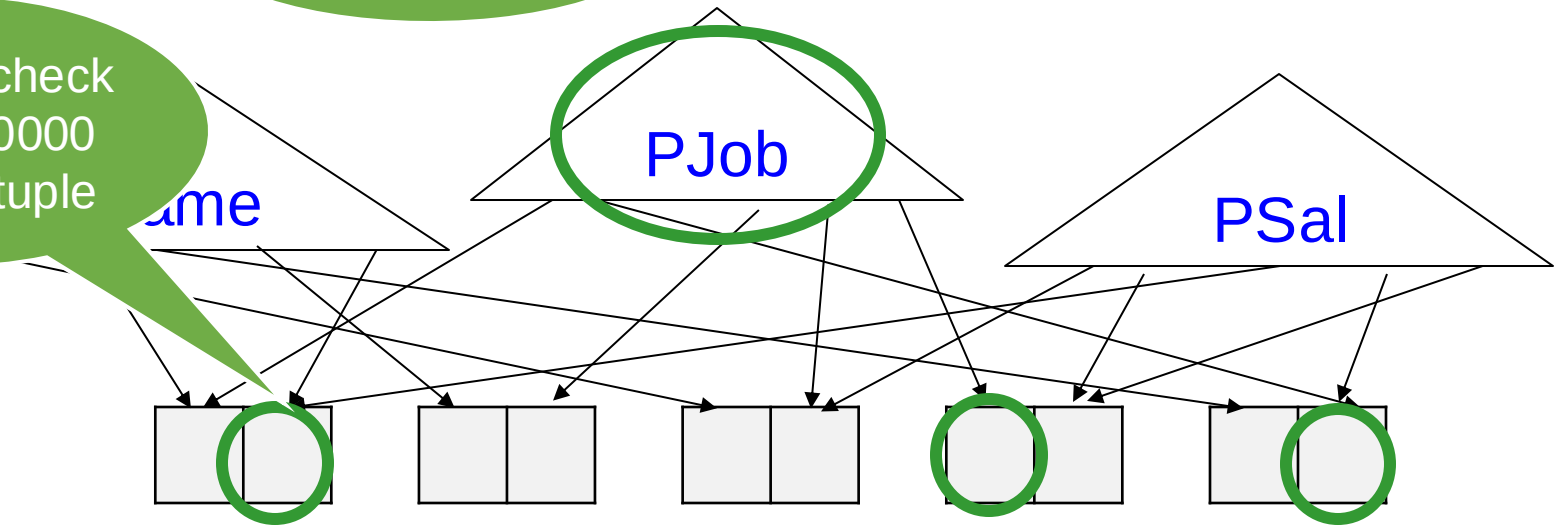
```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Access path:
use Pjob index

Must still check
salary=50000
for each tuple



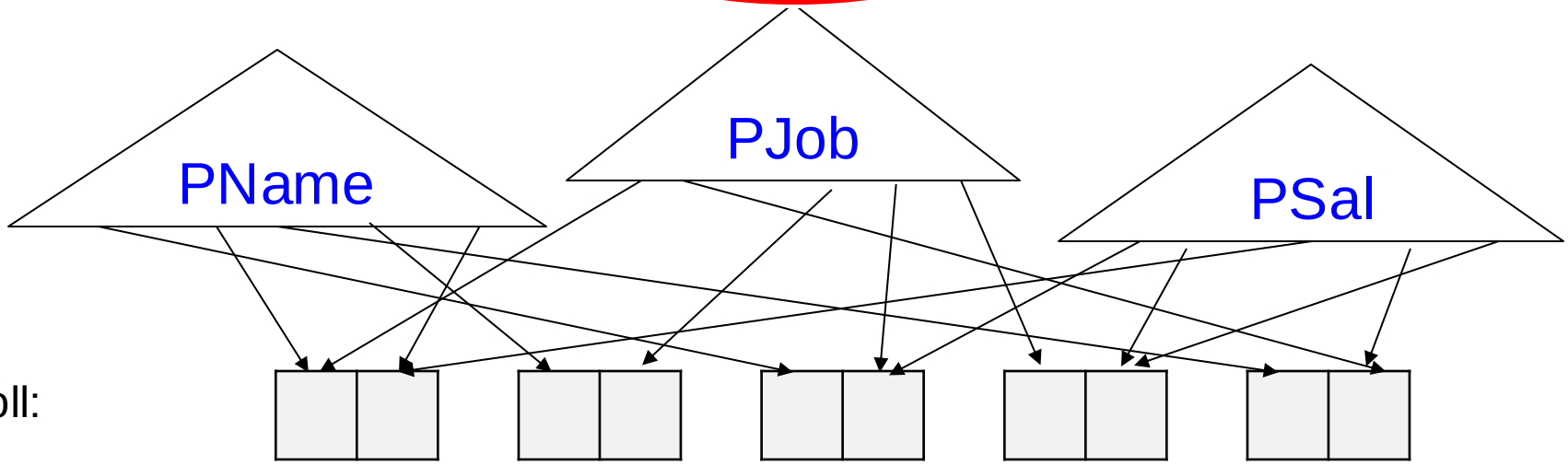
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Acces path:
use Psal index



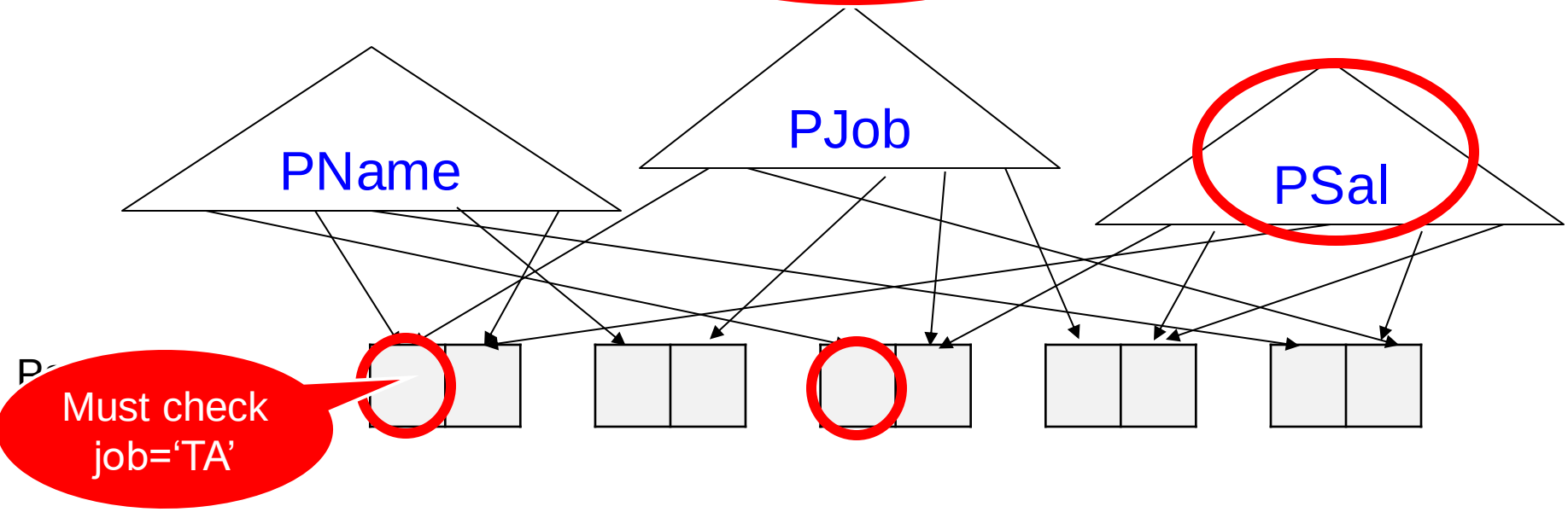
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Acces path:
use Psal index



Multi-attribute Index

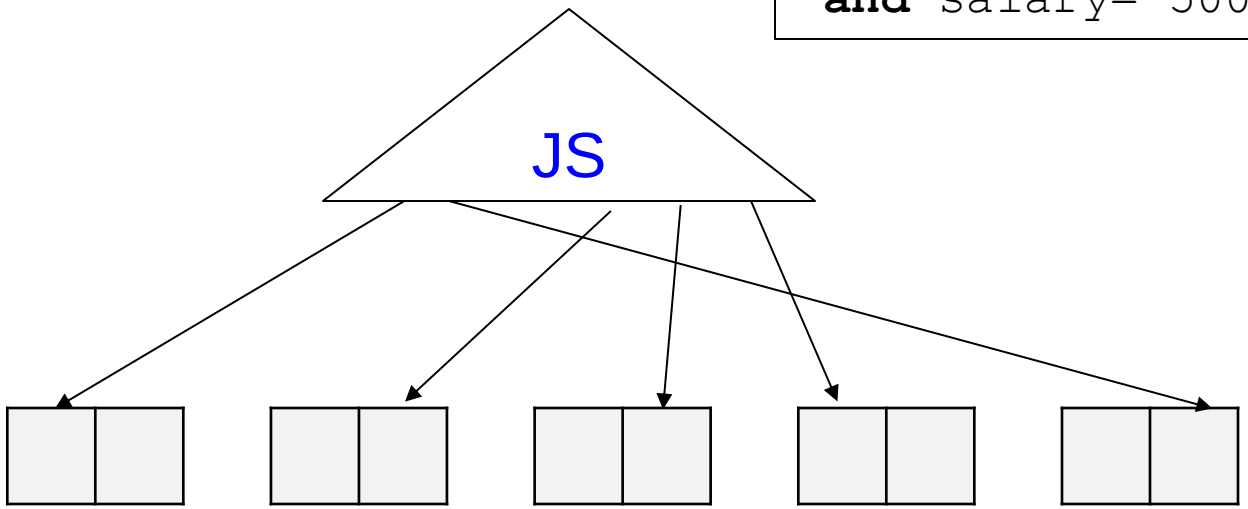
- An index can be on multiple attributes: A, B, C, …
  - Create index on R(A,B,C)
  - Create index on R(C,A,B)
- Values are concatenated, then sorted
- Attribute order matters

Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX JS on Payroll(job,salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary='50000';
```



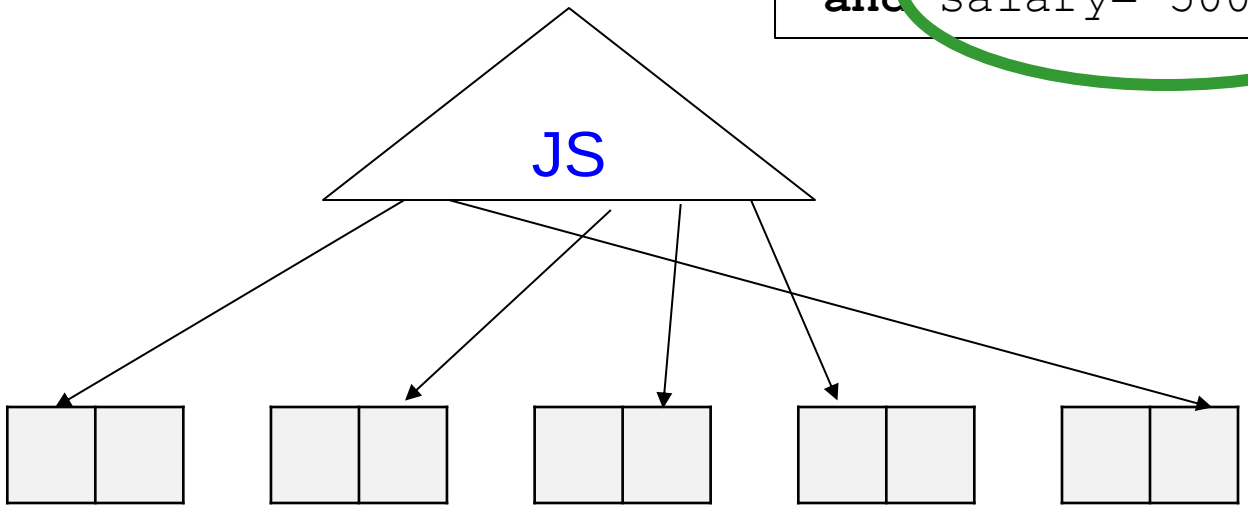
Payroll:

Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX JS on Payroll(job,salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
  and salary='50000';
```



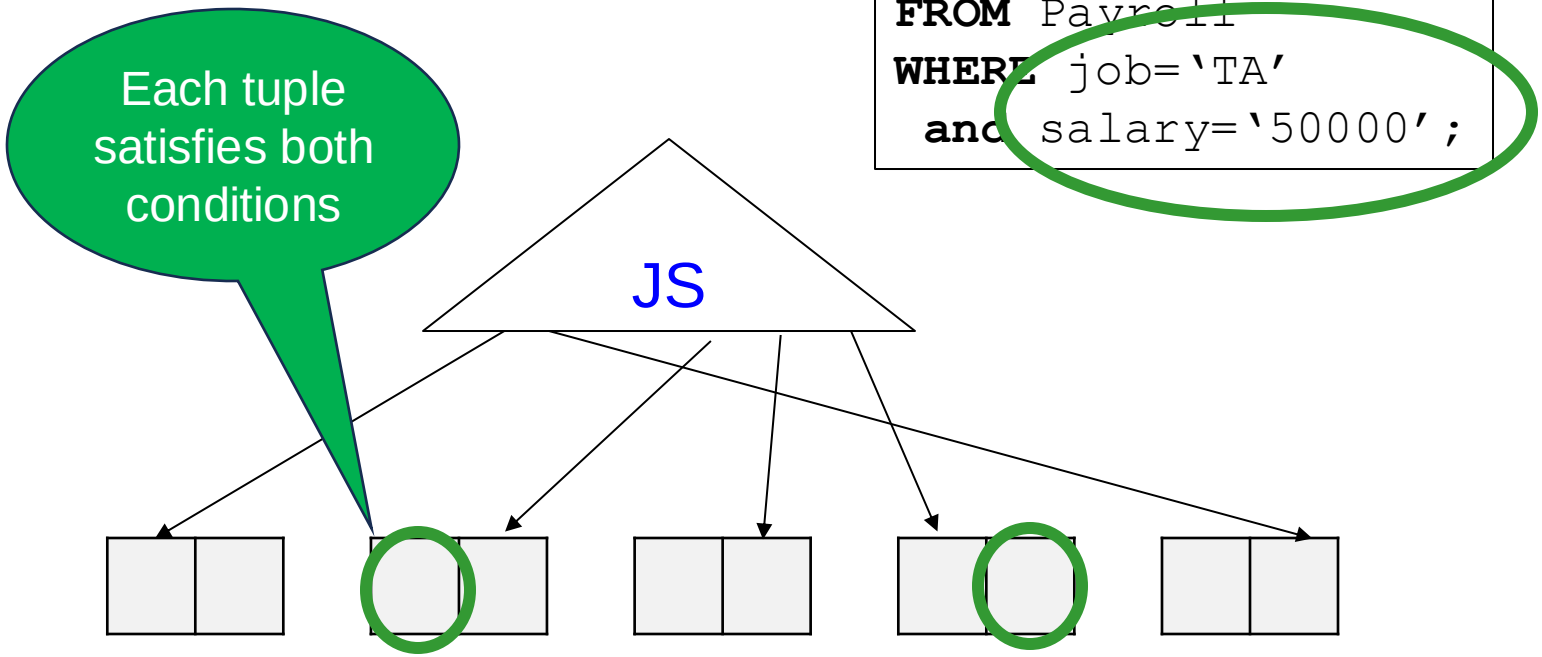
Payroll:

Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

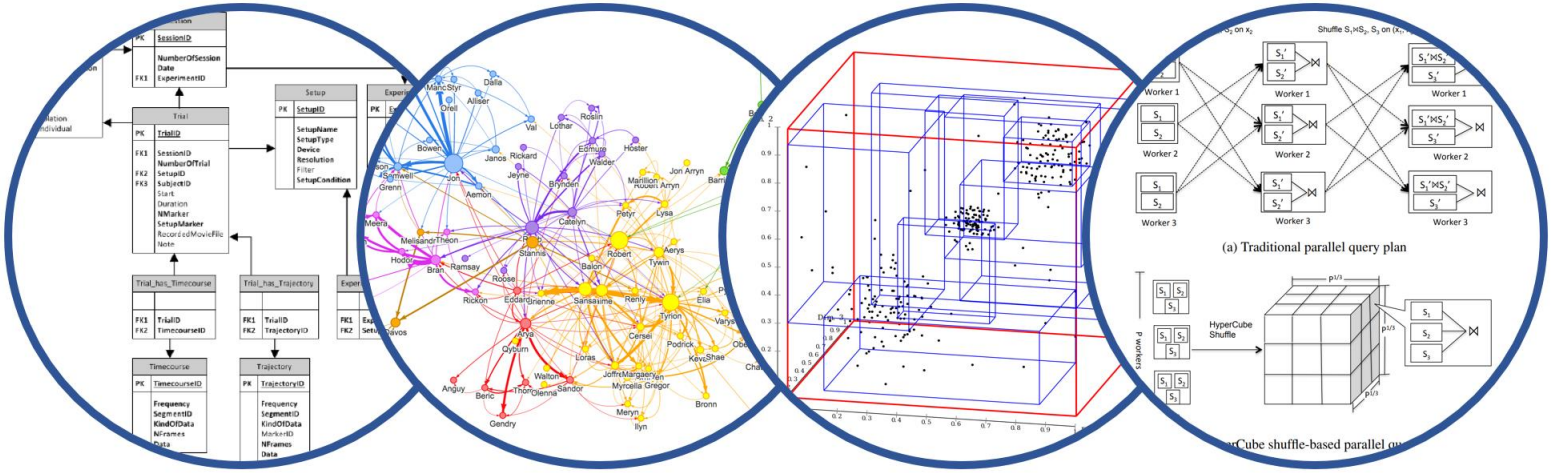
```
CREATE INDEX JS on Payroll(job,salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
  and salary='50000';
```



Multi-attribute Index

- A pair ordered lexicographically:
  - (‘Director’,200000) < (‘Prof’,50000)<(‘Prof’,200000)<…
- Only 1st attribute is known: range seach
  - Find (‘Prof’, *)
- Only 2nd attribute is known: impossible
  - Find (*, 200000)



Introduction to Data Management

Cardinality Estimation

Paul G. Allen School of Computer Science and Engineering
University of Washington, Seattle

Access Path Selection Wrapup

Multiple Indexes

- DB administrator can create multiple indexes
- But only one can be clustered…
- … and a selection can only use only one index

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

Payroll:

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

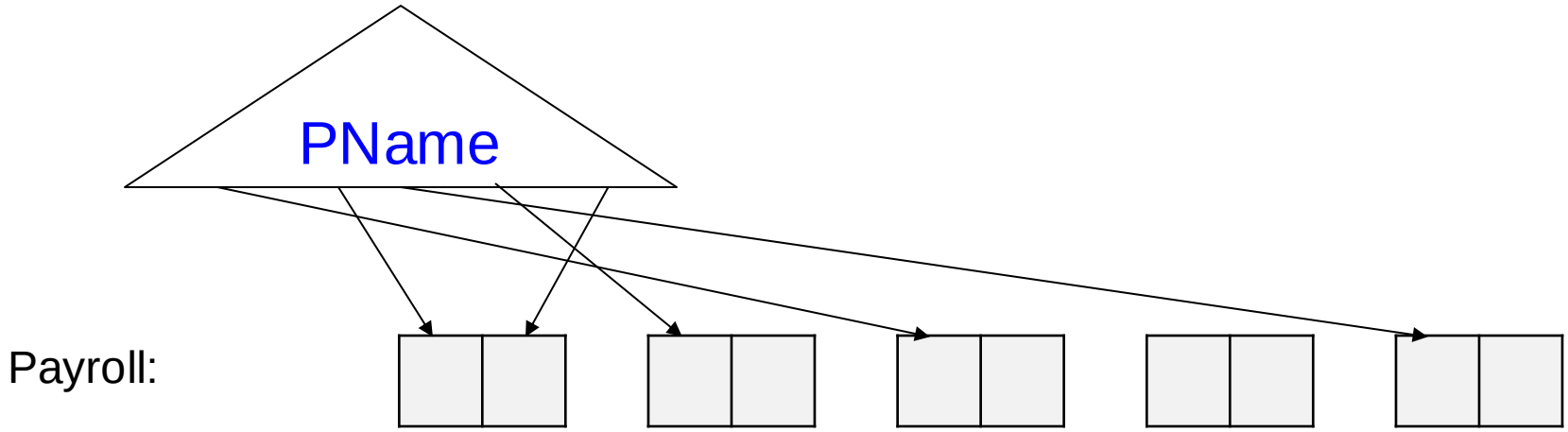
```
CREATE INDEX Pname on Payroll(name);
```

Payroll:

Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

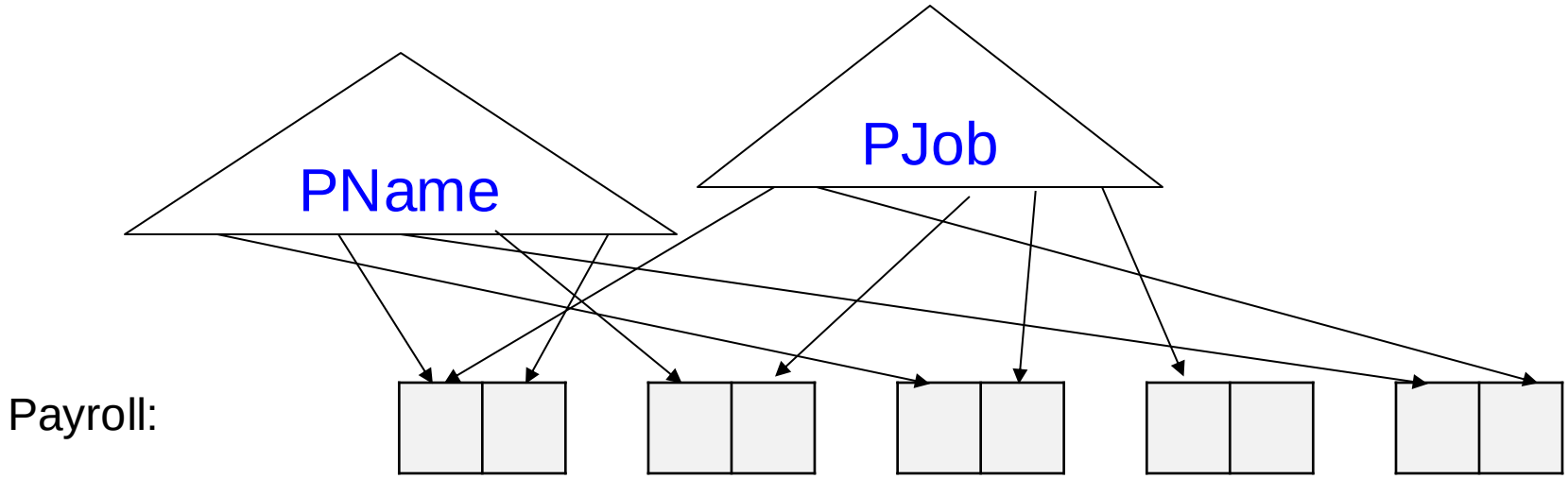
```
CREATE INDEX Pname on Payroll(name);
```



Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

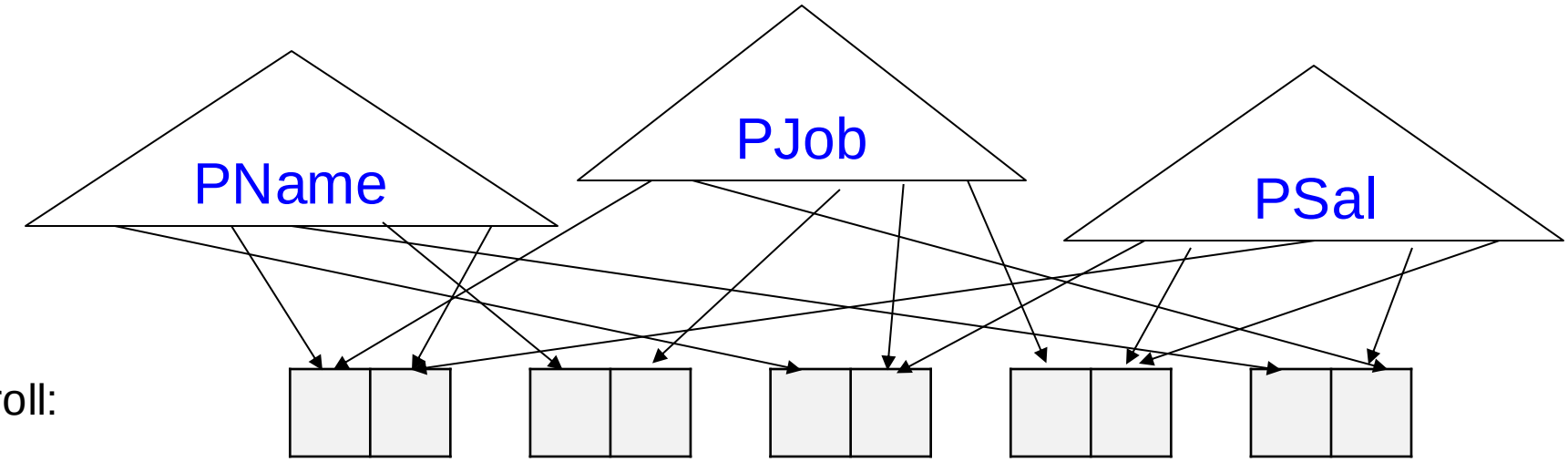
```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
```



Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

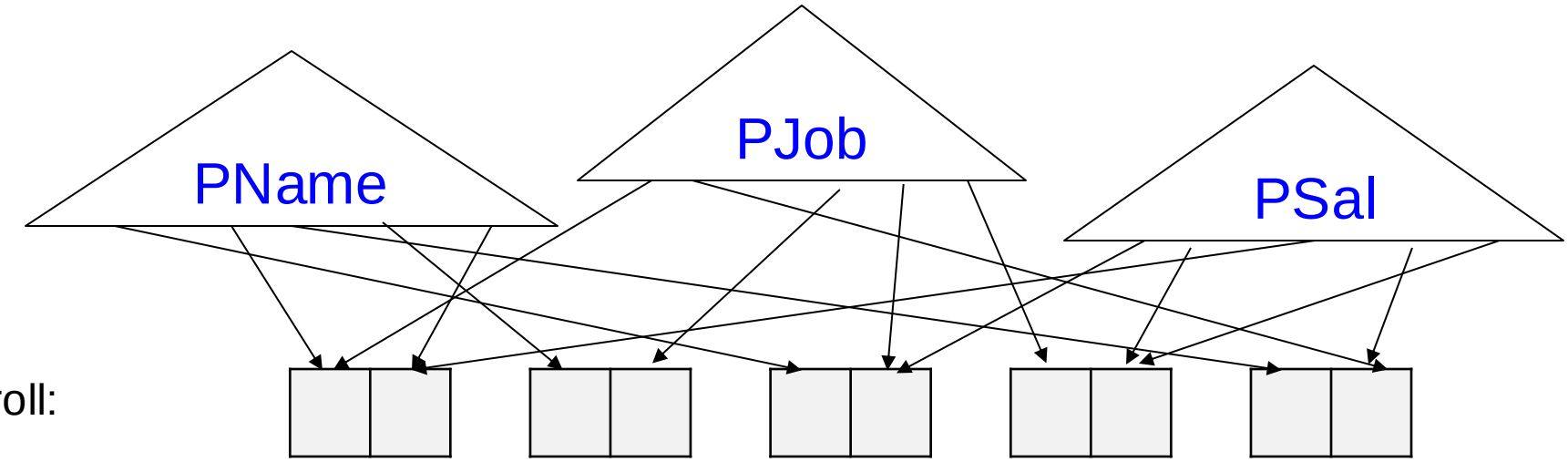


Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```



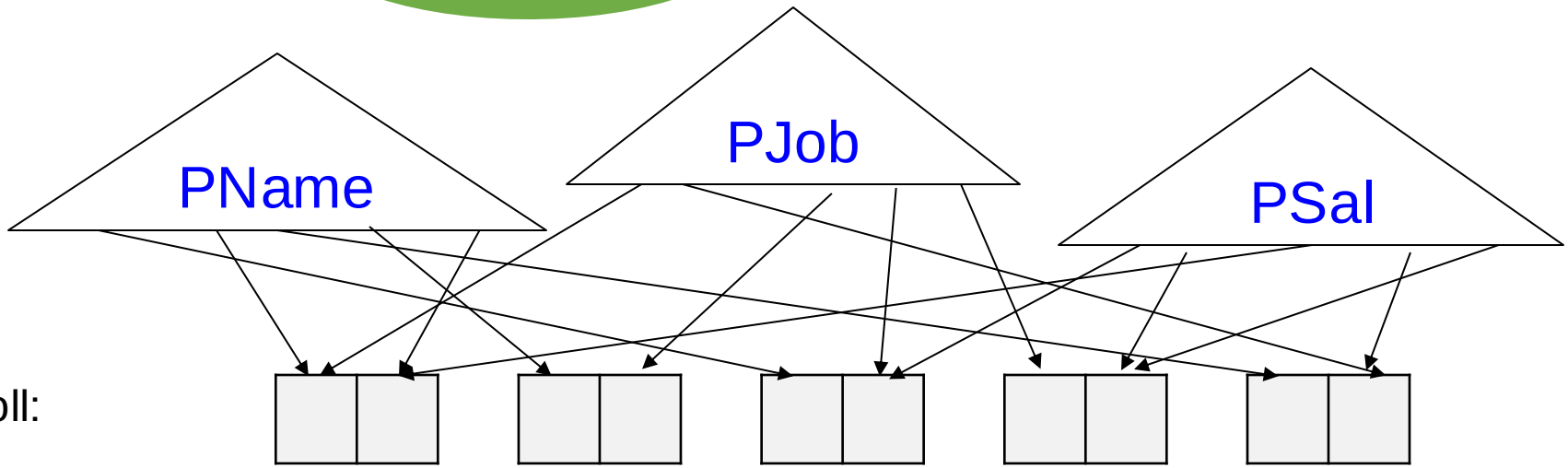
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Access path:
use Pjob index



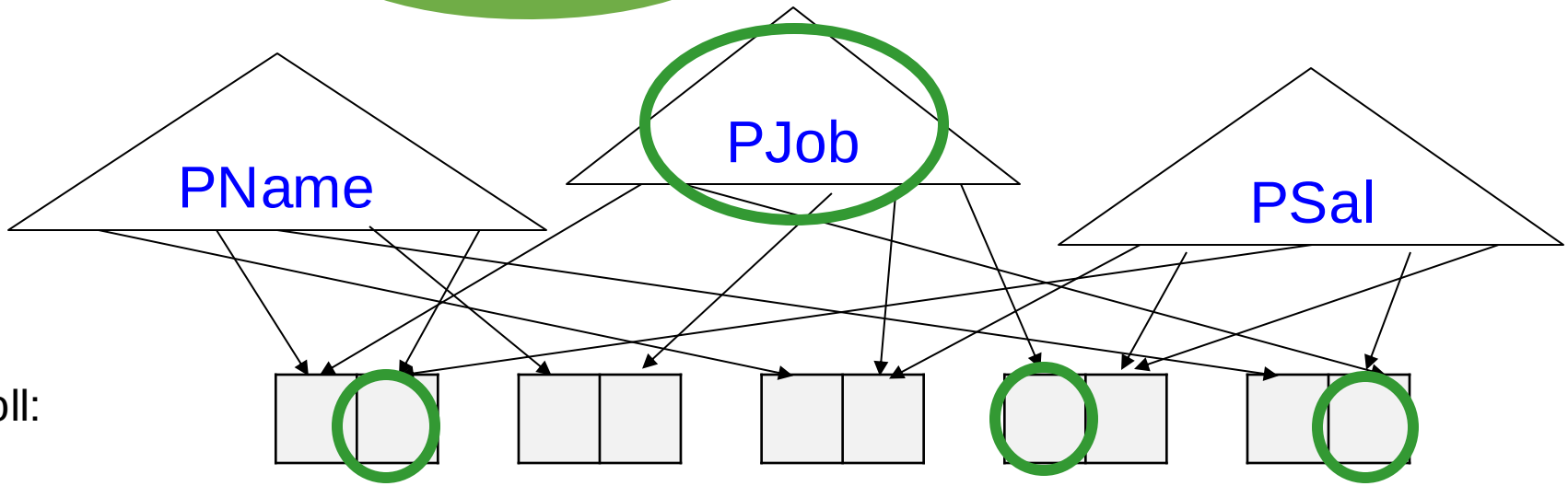
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Access path:
use Pjob index



Access Path Selection

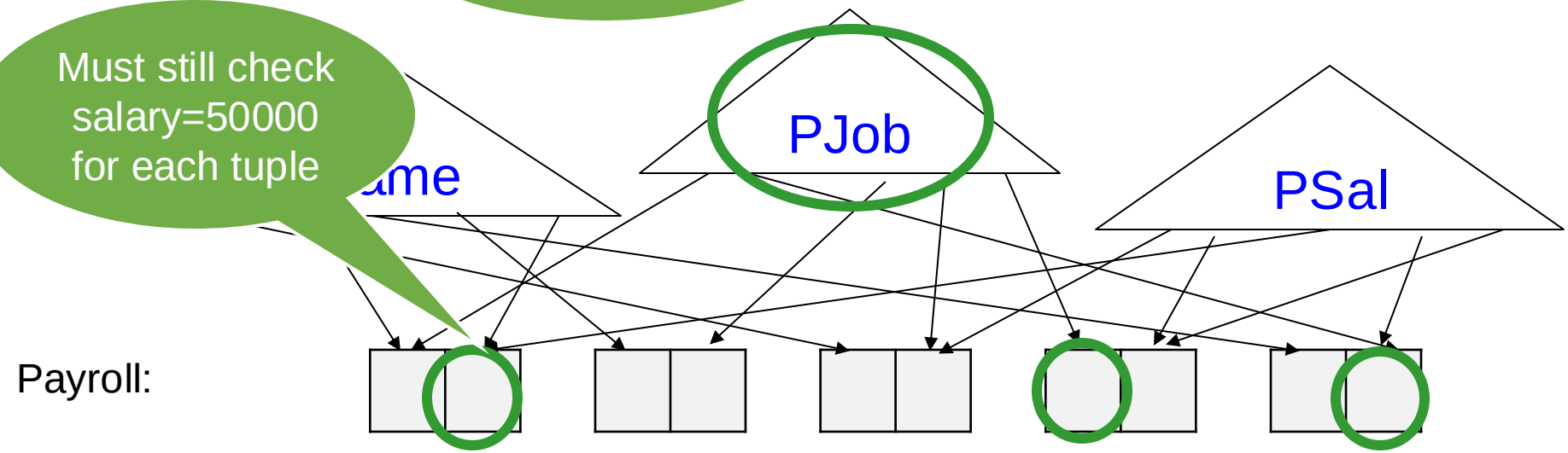
```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Access path:
use Pjob index

Must still check
salary=50000
for each tuple



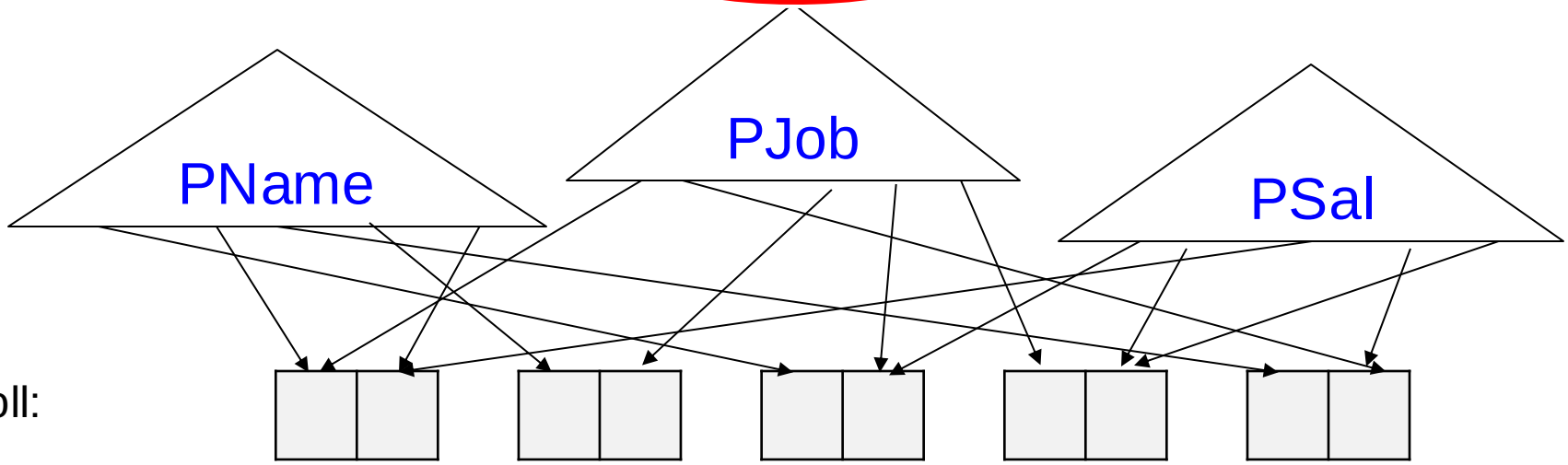
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Acces path:
use Psal index



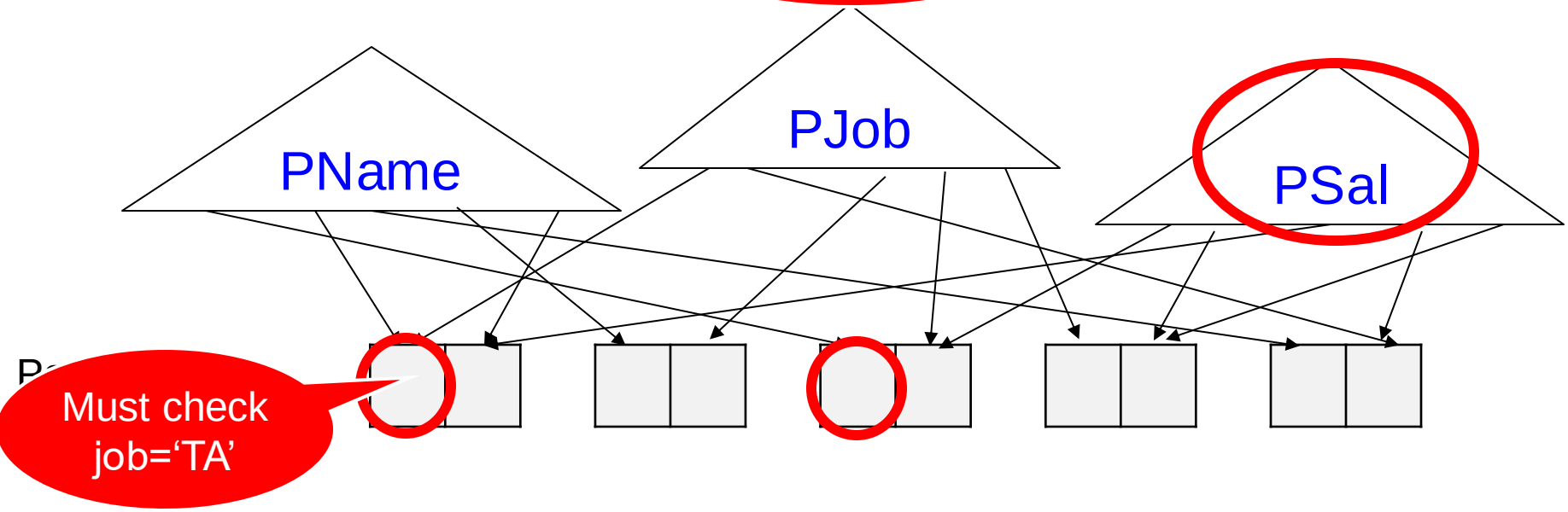
Access Path Selection

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX Pname on Payroll(name);
CREATE INDEX Pjob  on Payroll(job);
CREATE INDEX Psal  on Payroll(salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary=50000;
```

Acces path:
use Psal index



Multi-attribute Index

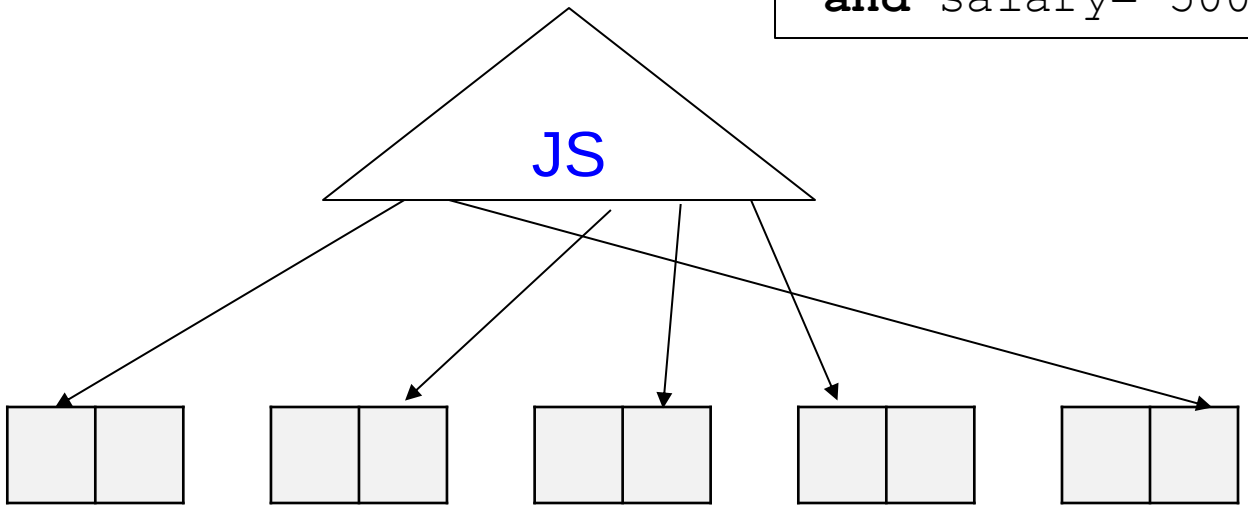
- An index can be on multiple attributes: A, B, C, …
  - Create index on R(B,A)
  - Create index on R(A,B,C)
  - Create index on R(C,A,B)
- Values are concatenated, then sorted
- Attribute order matters

Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX JS on Payroll(job,salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
 and salary='50000';
```



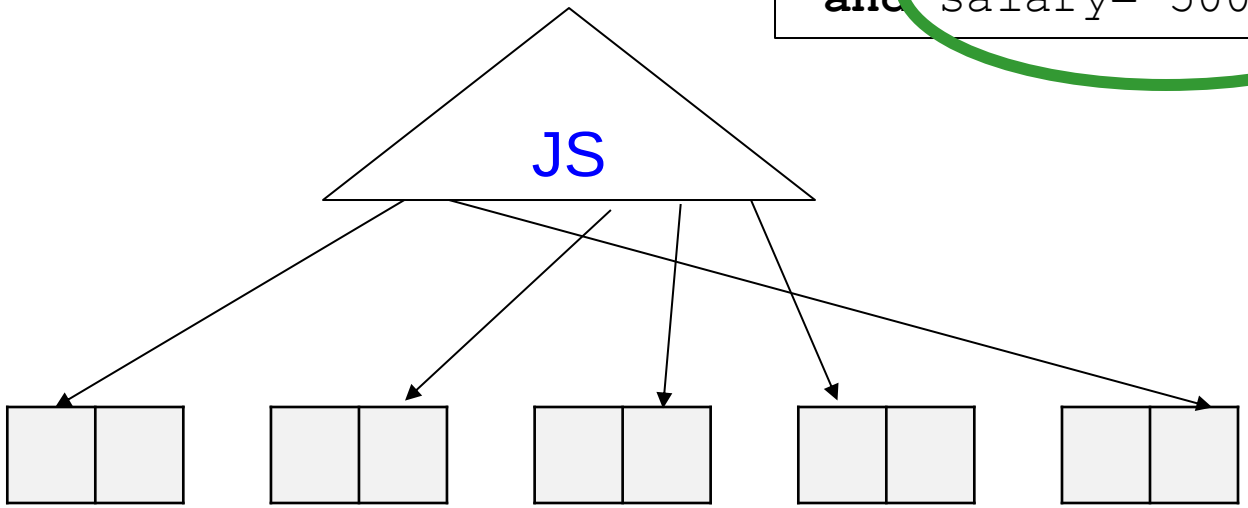
Payroll:

Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX JS on Payroll(job,salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
  and salary='50000';
```



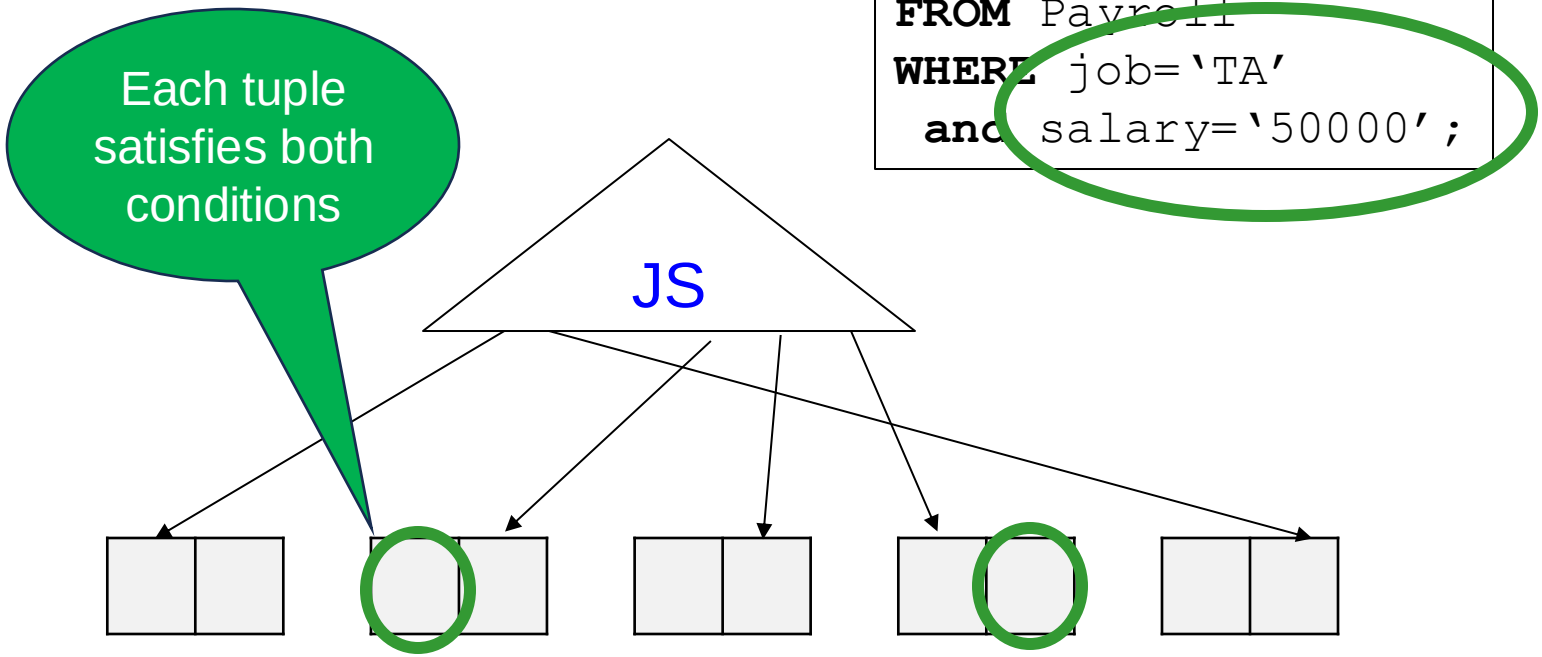
Payroll:

Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX JS on Payroll(job,salary);
```

```
SELECT *
FROM Payroll
WHERE job='TA'
  and salary='50000';
```



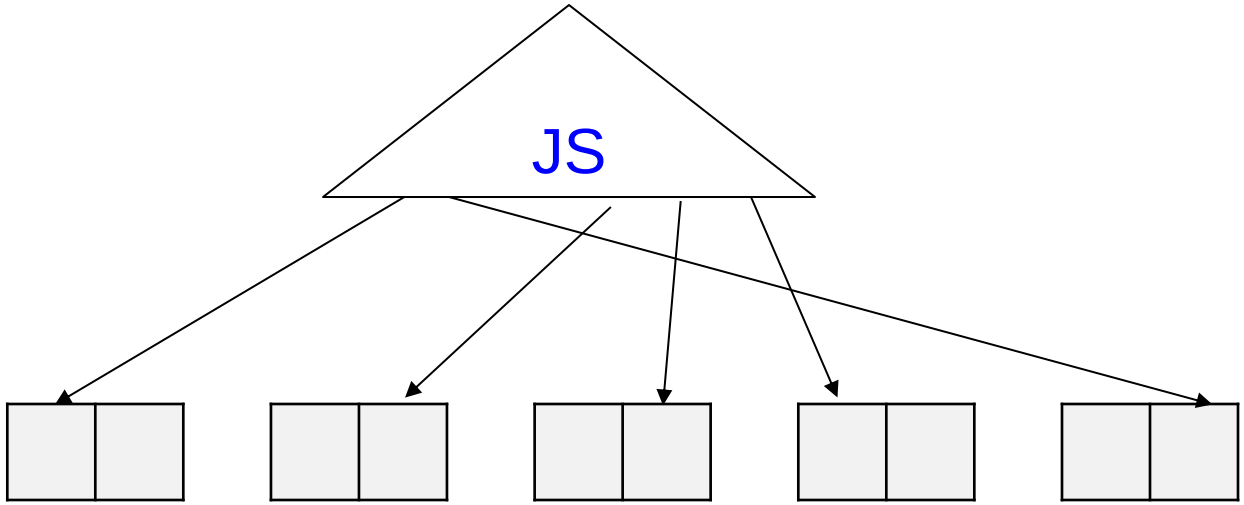
Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX JS on Payroll(job,salary);
```

Can we still use the index?

```
SELECT *
FROM Payroll
WHERE job='TA';
```



Payroll:

Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

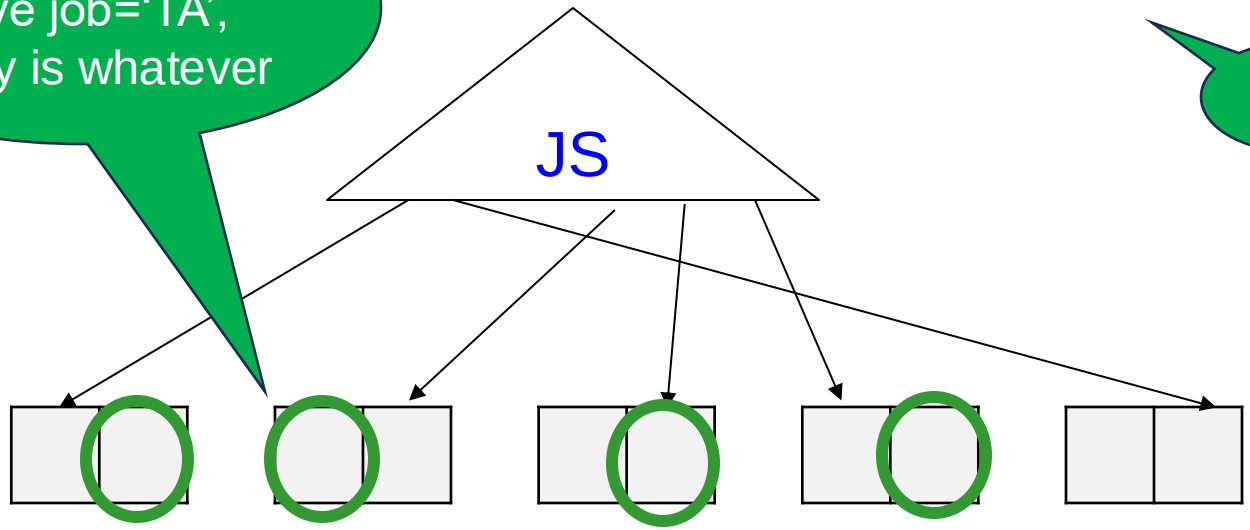
```
CREATE INDEX JS on Payroll(job,salary);
```

Can we still use the index?

```
SELECT *
FROM Payroll
WHERE job='TA';
```

YES

All these tuples have job='TA', salary is whatever



Payroll:

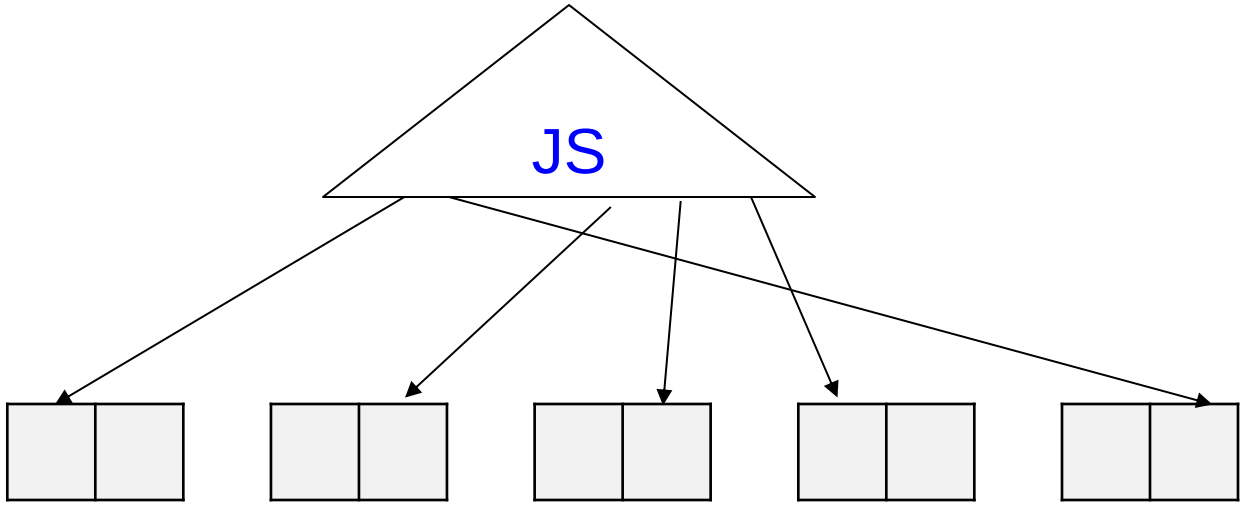
Multi-attribute Index

```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX JS on Payroll(job,salary);
```

Can we still use the index?

```
SELECT *
FROM Payroll
WHERE salary='50000';
```



Payroll:

Multi-attribute Index

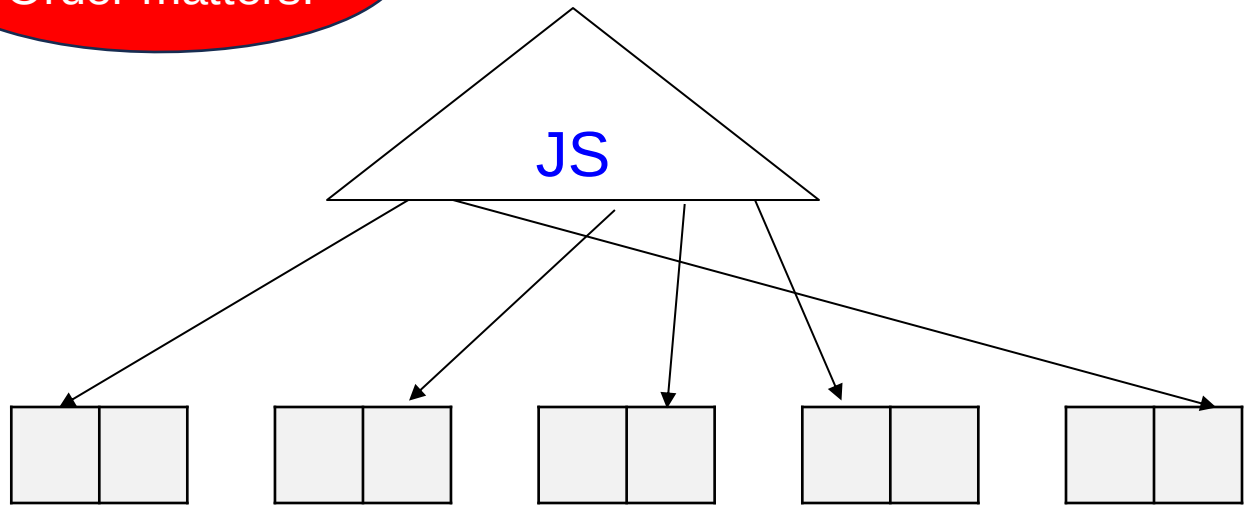
```
CREATE TABLE Payroll(UserID int, name text, job text, salary int)
```

```
CREATE INDEX JS on Payroll(job,salary);
```

NO
Order matters!

Can we still
use the index?

```
SELECT *
FROM Payroll
WHERE salary='50000';
```



Payroll:

Multi-attribute Index

- A pair ordered lexicographically:
  - (‘Director’,200000) < (‘Prof’,50000)<(‘Prof’,200000)<…
- Only 1st attribute is known: range seach
  - Find (‘Prof’, *)
- Only 2nd attribute is known: impossible
  - Find (*, 200000)

Cardinality Estimation

Overview

- DBMS keeps stats for each relation R(A,B,C,…):
  - Cardinality of R: T(R)
  - Number of distinct values in R.A: V(R,A)
    V(R, B), V(R, C), …
  - Histograms

- Cardinality estimation:
  Given only these stats, estimate output size.

Cardinality Estimation

Recursively on the query plan

Cardinality Estimation

Recursively on the query plan

- For a base table: Est(R) = T(R)

Cardinality Estimation

Recursively on the query plan

- For a base table: Est(R) = T(R)
- For an operator:
  $Est(\sigma_{pred}(R)) = \theta_{pred} * Est(R)$

Cardinality Estimation

Recursively on the query plan

- For a base table: Est(R) = T(R)

- For an operator:

$$Est(\sigma_{pred}(R)) = \theta_{pred} * Est(R)$$

$$Est(R \bowtie_{A=B} S) = \theta_{A=B} * Est(R) * Est(S)$$

Cardinality Estimation

Recursively on the query plan

- For a base table: Est(R) = T(R)

- For an operator:

$$Est(\sigma_{pred}(R)) = \theta_{pred} * Est(R)$$
$$Est(R \bowtie_{A=B} S) = \theta_{A=B} * Est(R) * Est(S)$$

θ is called the *selectivity factor*

Selectivity Factor

For $\sigma_{A=const}$ the selectivity factor is $\theta = \frac{1}{V(R,A)}$

Uniformity assumption

Selectivity Factor

For $\sigma_{A=const}$ the selectivity factor is $\theta = \frac{1}{V(R,A)}$

Uniformity assumption

T(Payroll) = 10000
V(Payroll, job) = 20

$\text{EST}\left[\sigma_{job='TA'}(\text{Payroll})\right] = ??$

Selectivity Factor

For $\sigma_{A=const}$ the selectivity factor is $\theta = \frac{1}{V(R,A)}$

Uniformity assumption

T(Payroll) = 10000
V(Payroll, job) = 20

$$EST\left[\sigma_{job='TA'}(Payroll)\right] = \frac{1}{20}10000 = 500$$

Selectivity Factor

For $\sigma_{p \text{ and } q}$ the sel. factor is $\theta_{p \text{ and } q} = \theta_p \times \theta_q$

Independence assumption

Selectivity Factor

For $\sigma_{p \text{ and } q}$ the sel. factor is $\theta_{p \text{ and } q} = \theta_p \times \theta_q$

Independence assumption

T(Payroll) = 10000
V(Payroll, job) = 20
V(Payroll, salary) = 50

$$\text{EST}\left[\sigma_{\text{job}=\text{'TA'} \text{ and } \text{salary}=20000}(\text{Payroll})\right] = ??$$

Selectivity Factor

For $\sigma_{p \text{ and } q}$ the sel. factor is $\theta_{p \text{ and } q} = \theta_p \times \theta_q$

Independence assumption

T(Payroll) = 10000
V(Payroll, job) = 20
V(Payroll, salary) = 50

$$\text{EST}\left[\sigma_{\text{job}=\text{'TA'} \text{ and } \text{salary}=20000}(\text{Payroll})\right] = \frac{1}{20}\frac{1}{50}10000 = 10$$

Selectivity Factor

For $\sigma_{p \text{ and } q}$ the sel. factor is $\theta_{p \text{ and } q} = \theta_p \times \theta_q$

Independence assumption

T(Payroll) = 10000
V(Payroll, job) = 20
V(Payroll, salary) = 50

This is likely an underestimate

$$\text{EST}\left[\sigma_{\text{job}=\text{'TA'} \text{ and } \text{salary}=20000}(\text{Payroll})\right] = \frac{1}{20}\frac{1}{50}10000 = 10$$

Selectivity Factor

$R \bowtie_{A=B} S$ the sel factor is $\theta = \frac{1}{\max(V(R,A),V(S,B))}$

Containment of values assumption

Selectivity Factor

$R \bowtie_{A=B} S$ the sel factor is $\theta = \frac{1}{\max(V(R,A),V(S,B))}$

Containment of values assumption

Justification:

- If $V(R,A) \leq V(S,B)$ then assume $R.A \subseteq S.B$
- 1 tuple in R joins $\frac{1}{V(S,B)} T(S)$ tuples in S
- Total output: $\frac{1}{V(S,B)} T(R)T(S)$

Selectivity Factor

$R \bowtie_{A=B} S$ the sel factor is $\theta = \frac{1}{\max(V(R,A),V(S,B))}$

Containment of values assumption

Justification:

- If $V(R,A) \leq V(S,B)$ then assume $R.A \subseteq S.B$
- 1 tuple in R joins $\frac{1}{V(S,B)} T(S)$ tuples in S
- Total output: $\frac{1}{V(S,B)} T(R)T(S)$

T(Payroll) = 10000
V(Payroll, UserID) = 10000

T(Regist) = 3000
V(Regist, UserID) = 2000

$EST[Payroll \bowtie Regist] =$??

Selectivity Factor

$R \bowtie_{A=B} S$ the sel factor is $\theta = \frac{1}{\max(V(R,A),V(S,B))}$

Containment of values assumption

Justification:

- If $V(R,A) \leq V(S,B)$ then assume $R.A \subseteq S.B$
- 1 tuple in R joins $\frac{1}{V(S,B)} T(S)$ tuples in S
- Total output: $\frac{1}{V(S,B)} T(R)T(S)$

T(Payroll) = 10000
V(Payroll, UserID) = 10000

T(Regist) = 3000
V(Regist, UserID) = 2000

$$EST[Payroll \bowtie Regist] = \frac{1}{\max(10000,2000)} 10000 \cdot 3000 = 3000$$

Summary of Assumptions

- Uniformity
- Independence
- Containment of values
- Preservation of values

Computing the Cost of a Plan

- Estimate cardinalities bottom-up
- Estimate cost by using estimated cardinalities

Histograms

- Histogram on R.A refines T(R), V(R,A)
- Each bucket contains: T(bucket), V(bucket, A)

Histograms

T(Payroll) = 10000, V(Payroll, salary) = 50

$$\mathrm{EST}\left[\sigma_{\mathrm{salary}=25\mathrm{k}}(\mathrm{Payroll})\right] = ???$$

Histograms

T(Payroll) = 10000, V(Payroll, salary) = 50

$$\text{EST}\left[\sigma_{\text{salary}=25\text{k}}(\text{Payroll})\right] = ??? \qquad = \frac{1}{50} 10000 = 10$$

Histograms

T(Payroll) = 10000, V(Payroll, salary) = 50

$$\text{EST}\left[\sigma_{\text{salary}=25\text{k}}(\text{Payroll})\right] = ??? \qquad = \frac{1}{50}10000 = 10$$

Salary:	0..20k	20k..29k	30k-39k	40k-49k	50k-59k	> 60k
T =	80	320	2000	4800	2600	200
V =	5	10	10	15	5	5

Histograms

T(Payroll) = 10000, V(Payroll, salary) = 50

$$\text{EST}\left[\sigma_{\text{salary}=25\text{k}}(\text{Payroll})\right] = ??? \qquad = \frac{1}{50}10000 = 10$$

Salary:	0..20k	20k..29k	30k-39k	40k-49k	50k-59k	> 60k
T =	80	320	2000	4800	2600	200
V =	5	10	10	15	5	5

Histograms

T(Payroll) = 10000, V(Payroll, salary) = 50

$$\text{EST}\left[\sigma_{\text{salary}=25\text{k}}(\text{Payroll})\right] = ??? \qquad = \frac{1}{50}10000 = 10$$

Salary:	0..20k	20k..29k	30k-39k	40k-49k	50k-59k	> 60k
T =	80	320	2000	4800	2600	200
V =	5	10	10	15	5	5

$$\text{EST}\left[\sigma_{\text{salary}=25\text{k}}(\text{Payroll})\right] = \frac{1}{10}320 = 32$$

Histograms

T(Payroll) = 10000, V(Payroll, salary) = 50

$$\text{EST}\left[\sigma_{\text{salary}=25\text{k}}(\text{Payroll})\right] = ??? \qquad = \frac{1}{50}10000 = 10$$

Salary:	0..20k	20k..29k	30k-39k	40k-49k	50k-59k	> 60k
T =	80	320	2000	4800	2600	200
V =	5	10	10	15	5	5

$$\text{EST}\left[\sigma_{\text{salary}=25\text{k}}(\text{Payroll})\right] = \frac{1}{10}320 = 32$$

A better estimate

Recap: How a Query Engine Works

- Each RA operator has multiple physical operators
- Indexes: speed up some queries slow down others
- Rewrite rules transform one query plan to another
- Cardinality estimators used to estimate the cost; they are often wrong

Real Database Engines

Leis et al., *How Good Are Query Optimizers, Really?*, VLDB 2015

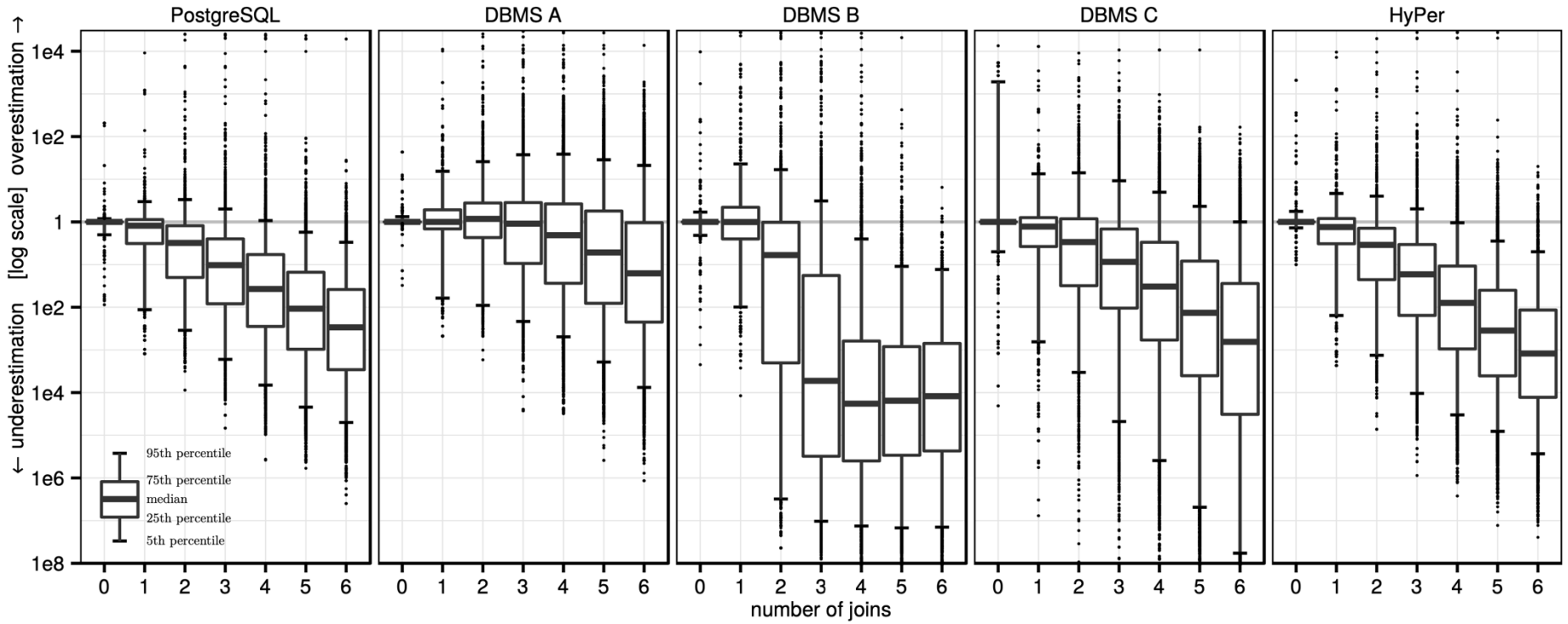


Figure 3: Quality of cardinality estimates for multi-join queries in comparison with the true cardinalities. Each boxplot summarizes the error distribution of all subexpressions with a particular size (over all queries in the workload)

Real Database Engines

Leis et al., *How Good Are Query Optimizers, Really?*, VLDB 2015

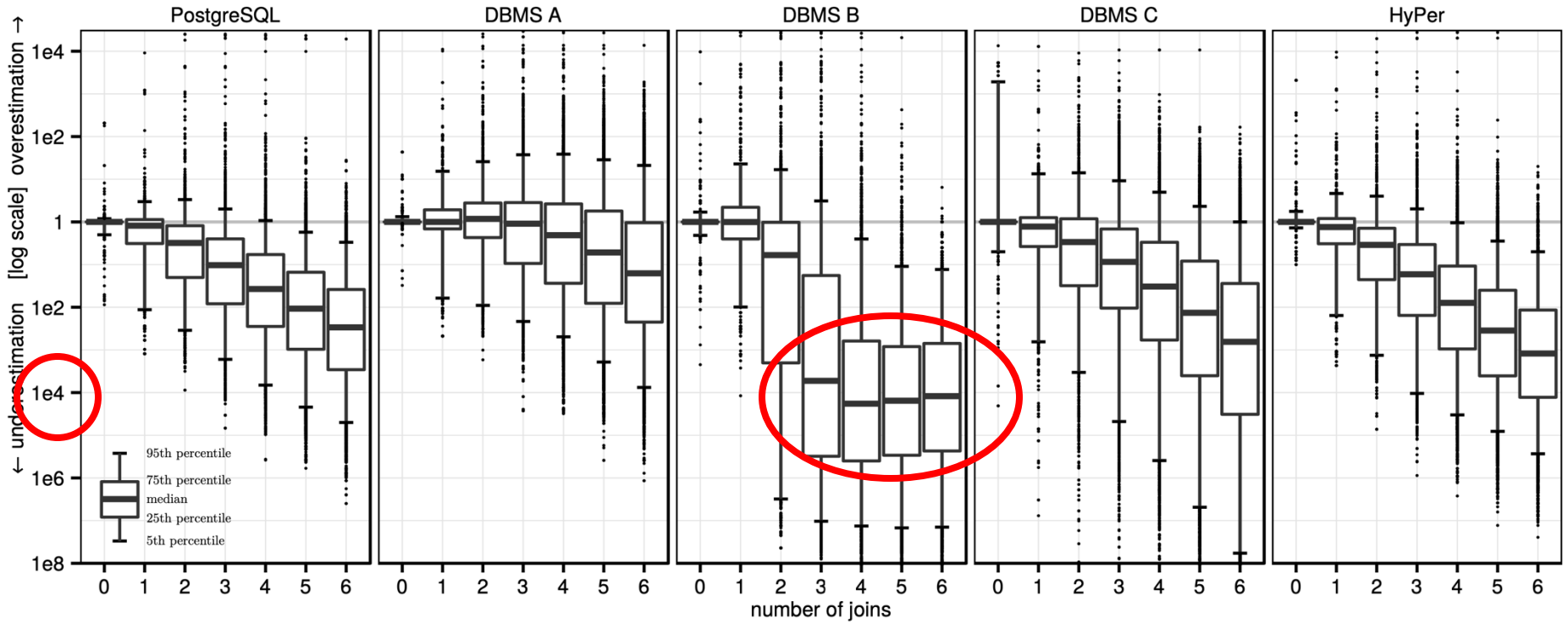


Figure 3: Quality of cardinality estimates for multi-join queries in comparison with the true cardinalities. Each boxplot summarizes the error distribution of all subexpressions with a particular size (over all queries in the workload)