

Query Processing

P.J. McBrien

Imperial College London

Rows and Pages

account=[$a_{100}, a_{101}, a_{103}, a_{107}, a_{119}, a_{125}$]

movement=[$m_{1000}, m_{1001}, m_{1002}, m_{1004}, m_{1005}, m_{1006}, m_{1007}, m_{1008}, m_{1009}$]

Pages(account) = [P_1, P_2, P_3]

$P_1 = [a_{100}, a_{101}]$, $P_2 = [a_{103}, a_{107}]$, $P_3 = [a_{119}, a_{125}]$

Pages(movement) = [P_4, P_5, P_6, P_7]

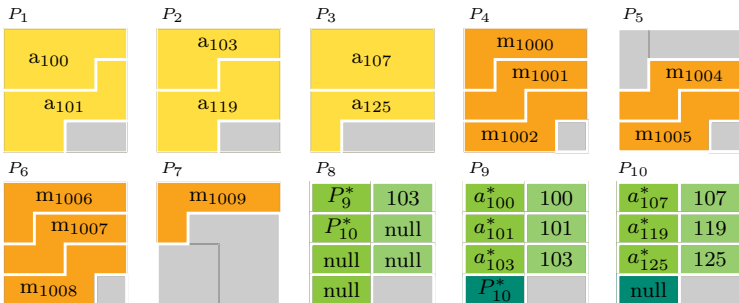
$P_4 = [m_{1000}, m_{1001}, m_{1002}]$, $P_5 = [m_{1004}, m_{1005}]$, $P_6 = [m_{1006}, m_{1007}, m_{1008}]$, $P_7 = [m_{1009}]$

Pages(index(account.no)) = [P_8, P_9, P_{10}]

$P_8 = [P_9^*, 103, P_{10}^*, null, null, null, null]$, $P_9 = [a_{100}^*, 100, a_{101}^*, 101, a_{103}^*, 103, P_{10}^*]$,

$P_{10} = [a_{107}^*, 107, a_{119}^*, 119, a_{125}^*, 125, null]$

Page Layout on Disc



Rows and Pages

File metrics

$|account| = 6$

$|movement| = 9$

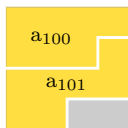
$|Pages(account)| = 3$

$|Pages(movement)| = 4$

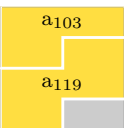
$|Pages(index(account.no))| = 3$

Page Layout on Disc

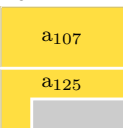
P_1



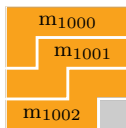
P_2



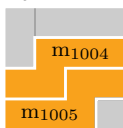
P_3



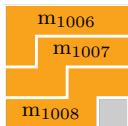
P_4



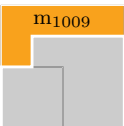
P_5



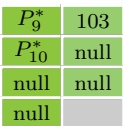
P_6



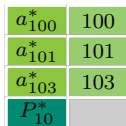
P_7



P_8



P_9



P_{10}



Quiz 1: Pages in a B⁺-tree

A B⁺-tree has fan out of 5, and the table it indexes has 80 rows.

What is the minimum number of pages occupied by the B⁺-tree?

A

21

B

25

C

26

D

31

Example Query

account				
no	type	cname	rate	sortcode
100	'current'	'McBrien, P.'	NULL	67
101	'deposit'	'McBrien, P.'	5.25	67
103	'current'	'Boyd, M.'	NULL	34
107	'current'	'Poulovassilis, A.'	NULL	56
119	'deposit'	'Poulovassilis, A.'	5.50	56
125	'current'	'Bailey, J.'	NULL	56

```

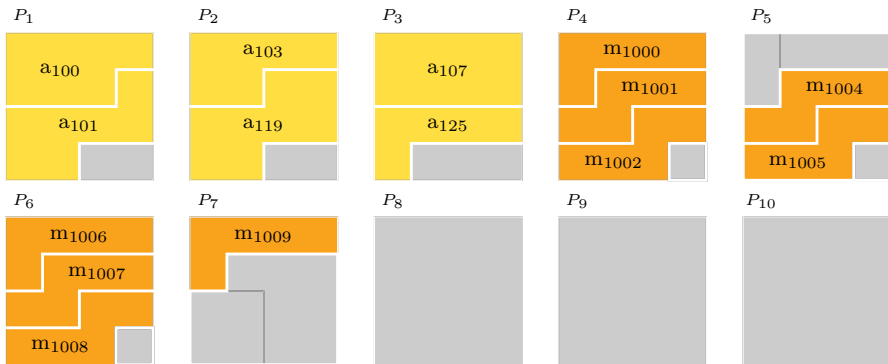
SELECT  movement.mid
        account.no,
        account.cname,
FROM    account JOIN movement
        ON account.no=movement.no
  
```



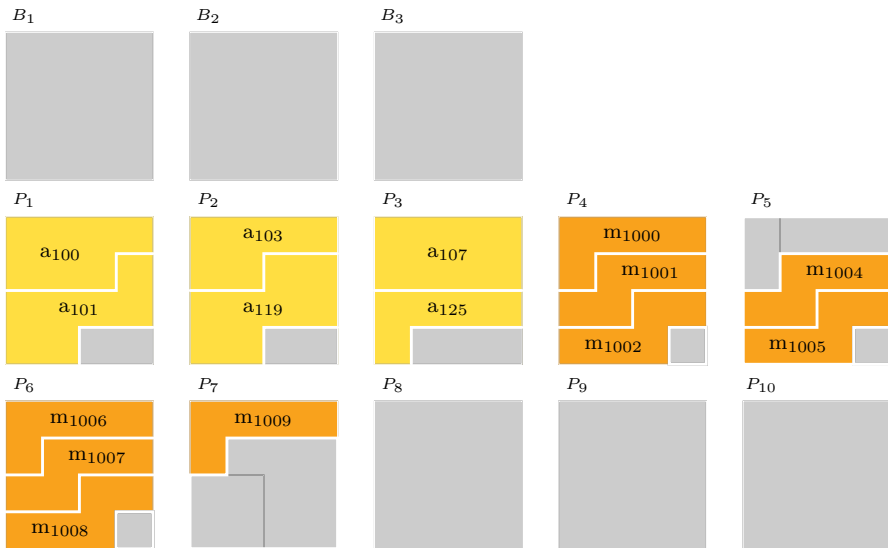
movement			
mid	no	amount	tdate
1000	100	2300.00	5/1/1999
1001	101	4000.00	5/1/1999
1002	100	-223.45	8/1/1999
1004	107	-100.00	11/1/1999
1005	103	145.50	12/1/1999
1006	100	10.23	15/1/1999
1007	107	345.56	15/1/1999
1008	101	1230.00	15/1/1999
1009	119	5600.00	18/1/1999

query1		
mid	no	cname
1000	100	'McBrien, P.'
1001	101	'McBrien, P.'
1002	100	'McBrien, P.'
1006	100	'McBrien, P.'
1008	101	'McBrien, P.'
1005	103	'Boyd, M.'
1009	119	'Poulovassilis, A.'
1004	107	'Poulovassilis, A.'
1007	107	'Poulovassilis, A.'

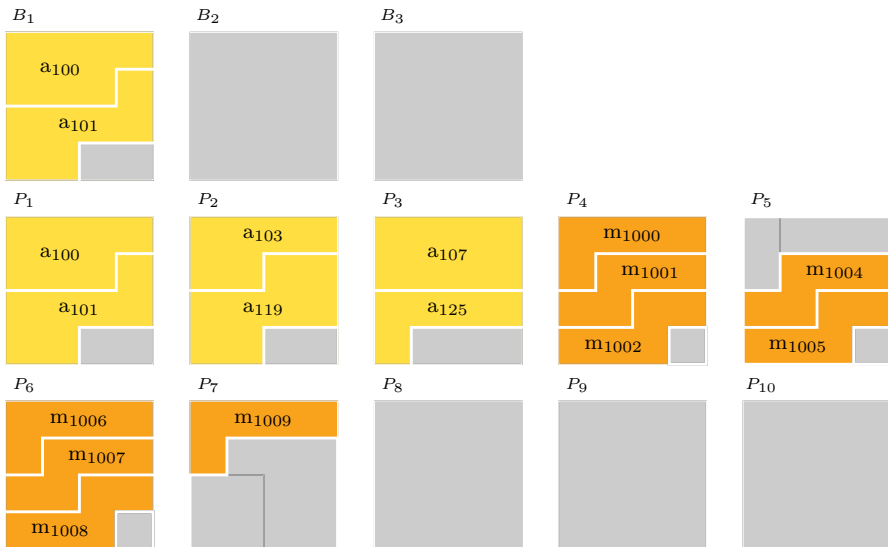
Block Nested Loop Join: Minimum Buffer Space



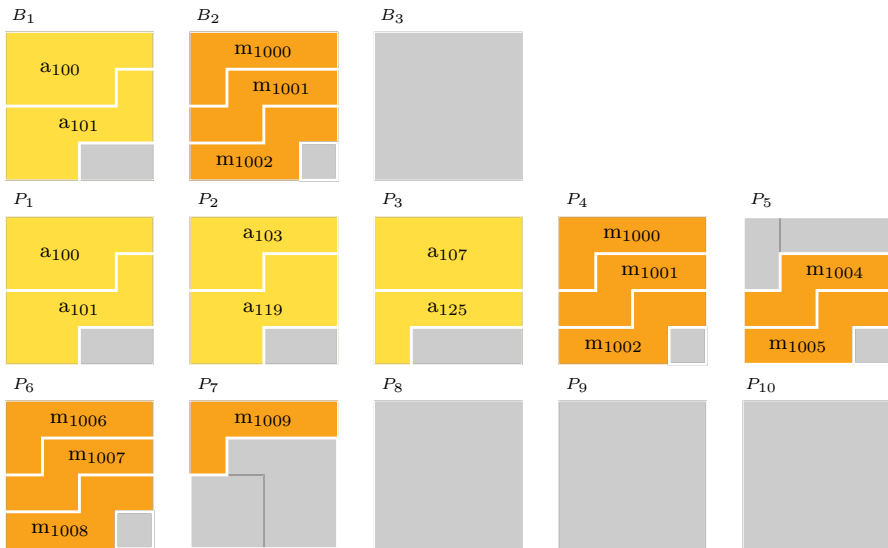
Block Nested Loop Join: Minimum Buffer Space



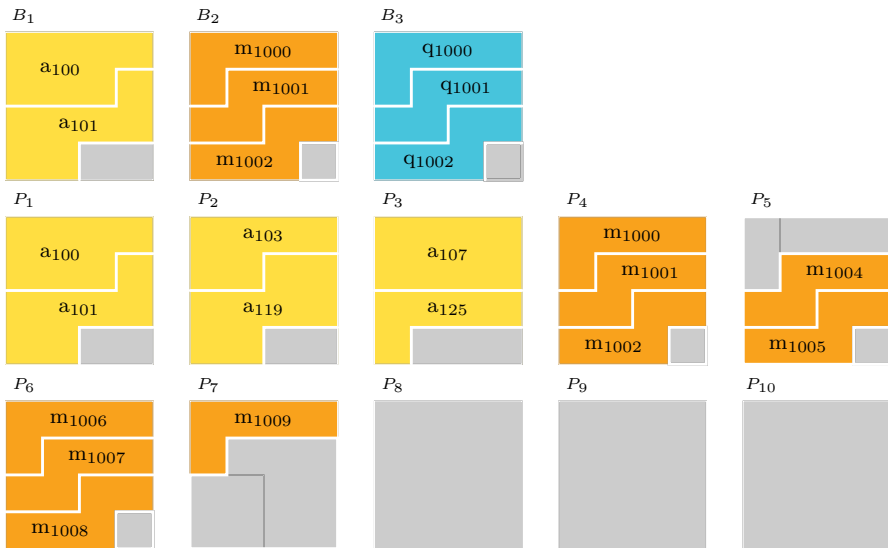
Block Nested Loop Join: Minimum Buffer Space



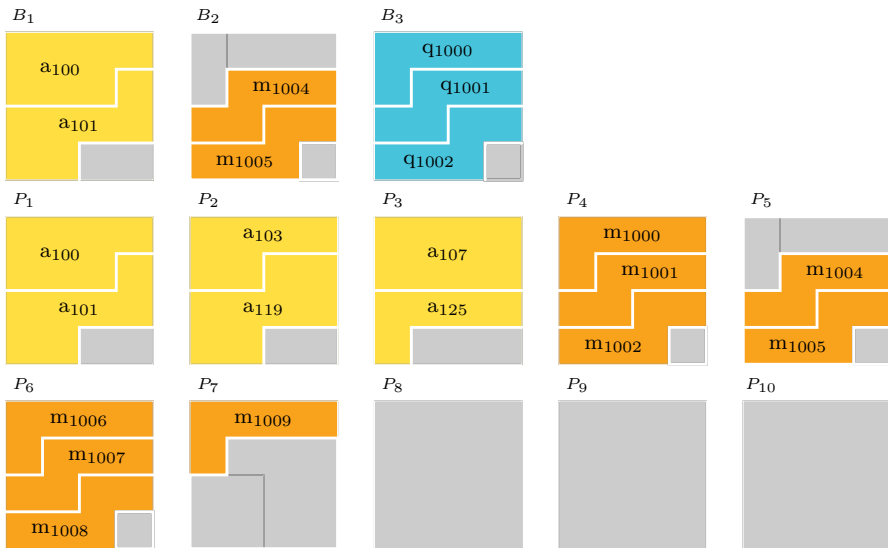
Block Nested Loop Join: Minimum Buffer Space



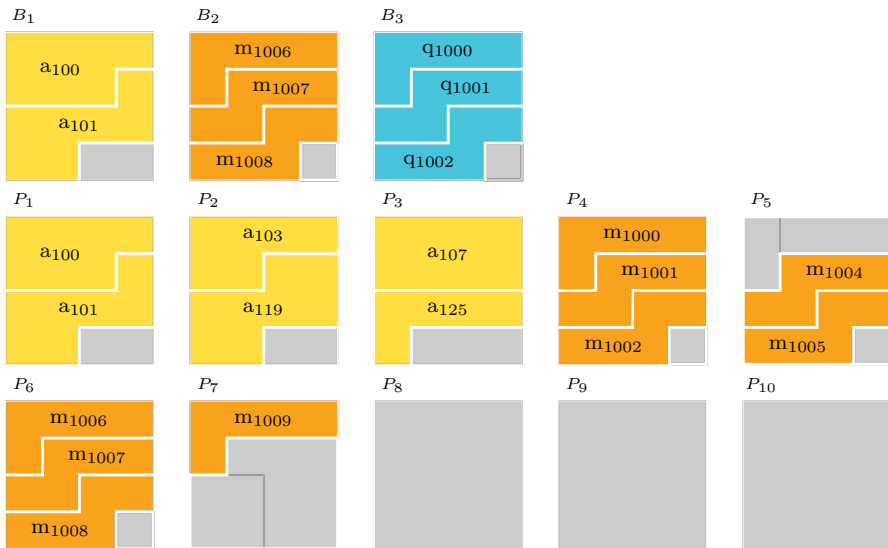
Block Nested Loop Join: Minimum Buffer Space



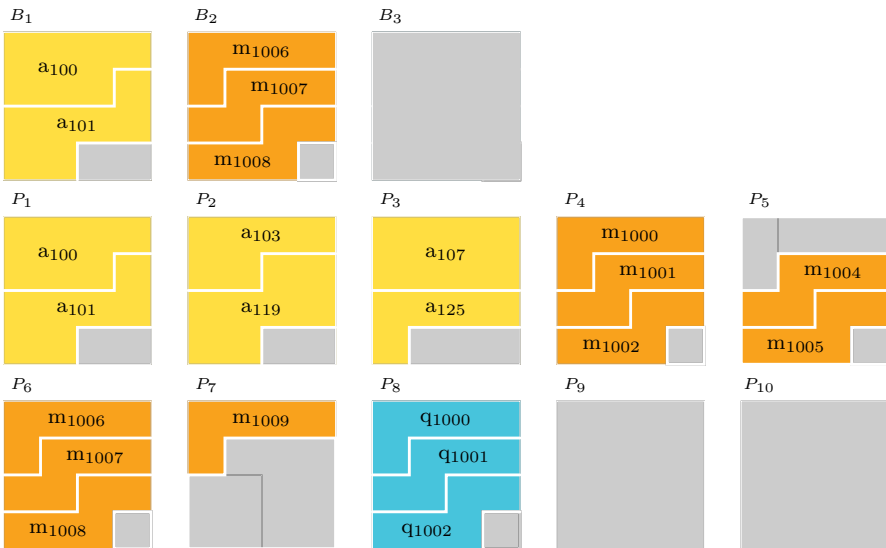
Block Nested Loop Join: Minimum Buffer Space



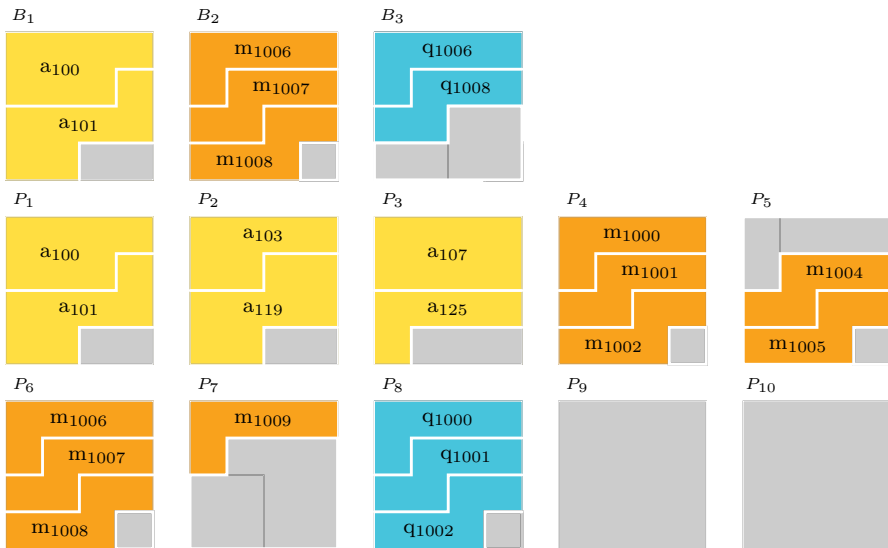
Block Nested Loop Join: Minimum Buffer Space



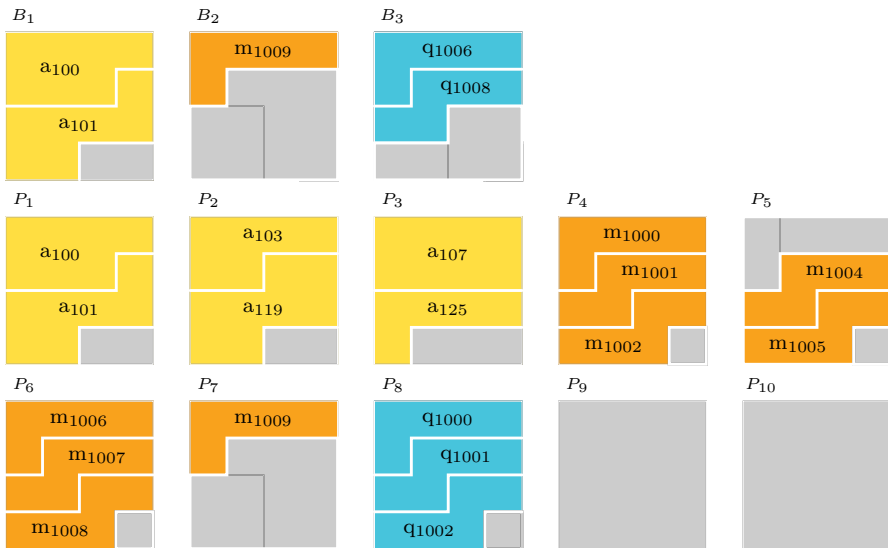
Block Nested Loop Join: Minimum Buffer Space



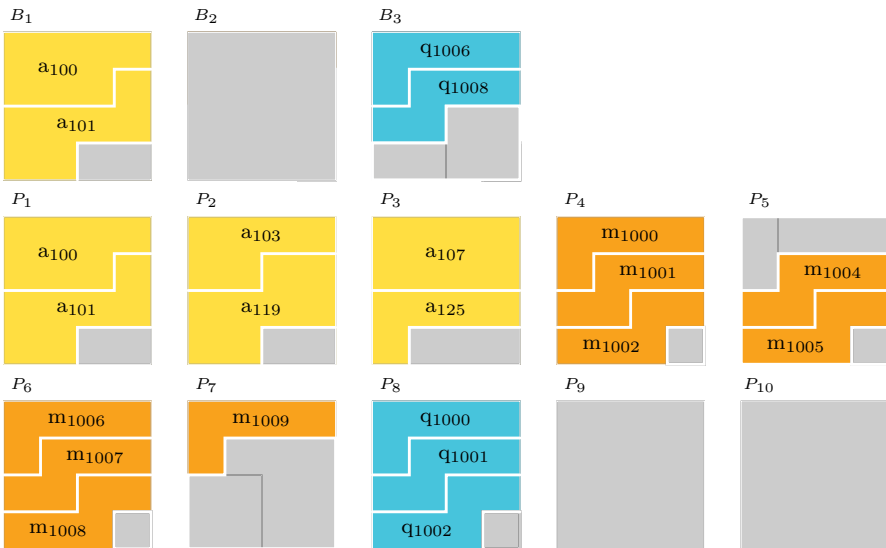
Block Nested Loop Join: Minimum Buffer Space



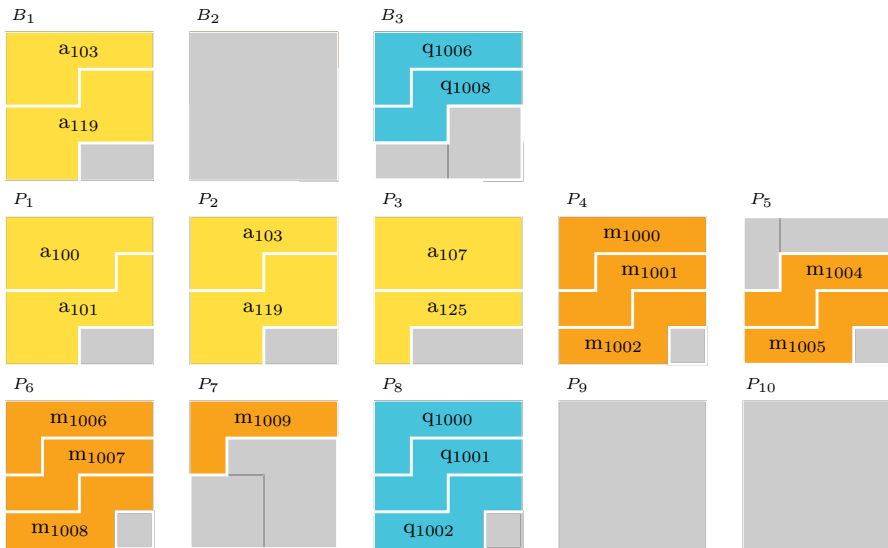
Block Nested Loop Join: Minimum Buffer Space



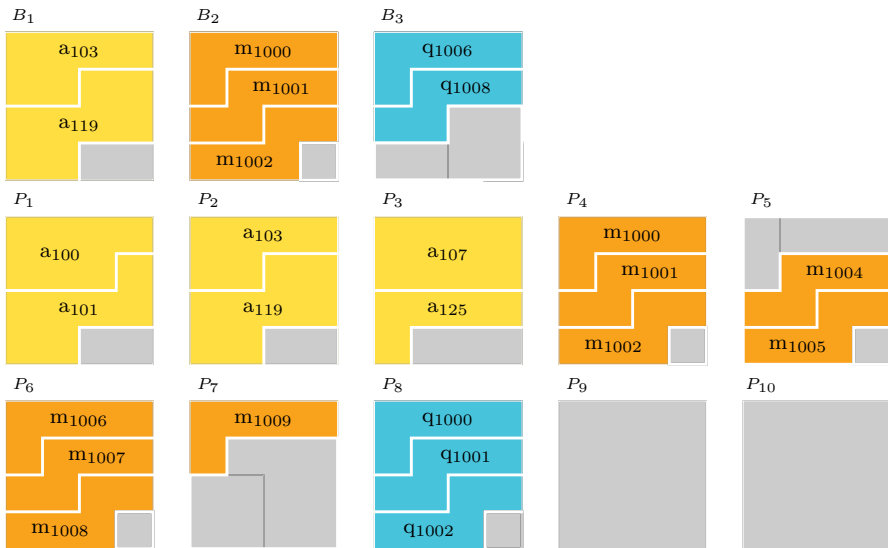
Block Nested Loop Join: Minimum Buffer Space



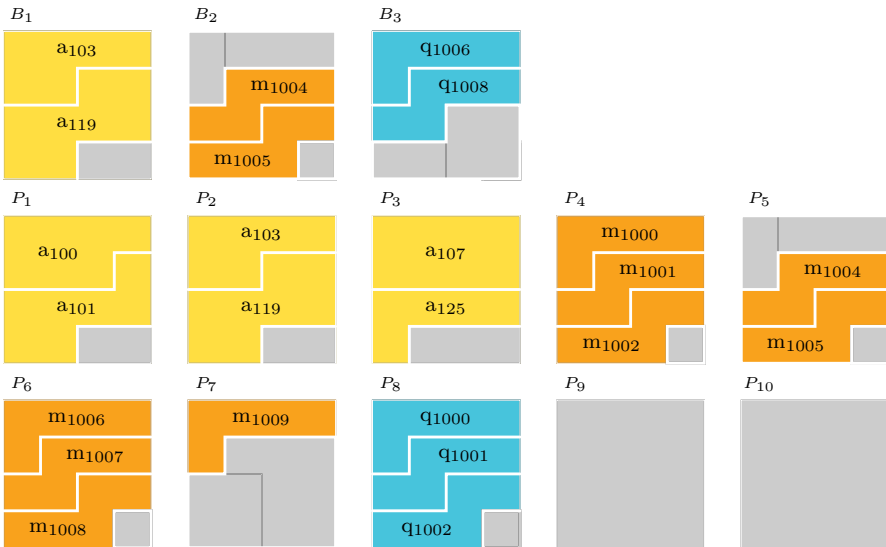
Block Nested Loop Join: Minimum Buffer Space



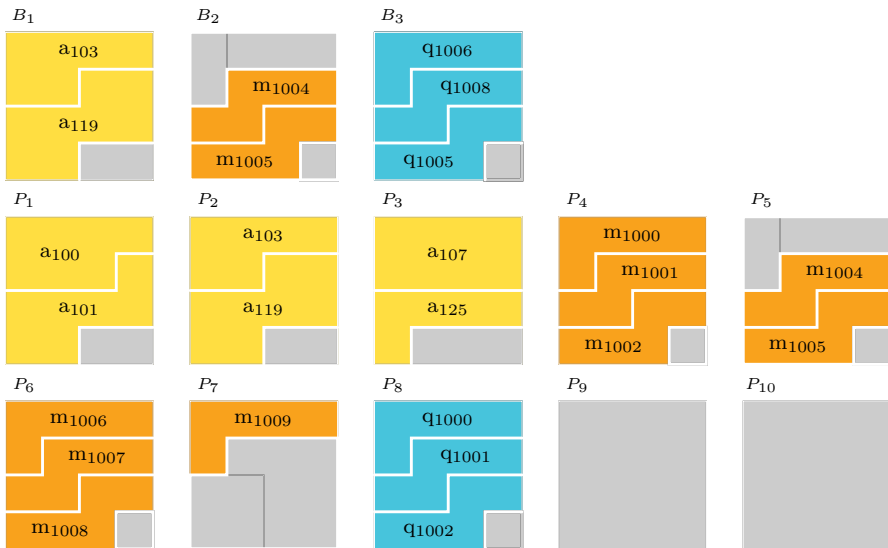
Block Nested Loop Join: Minimum Buffer Space



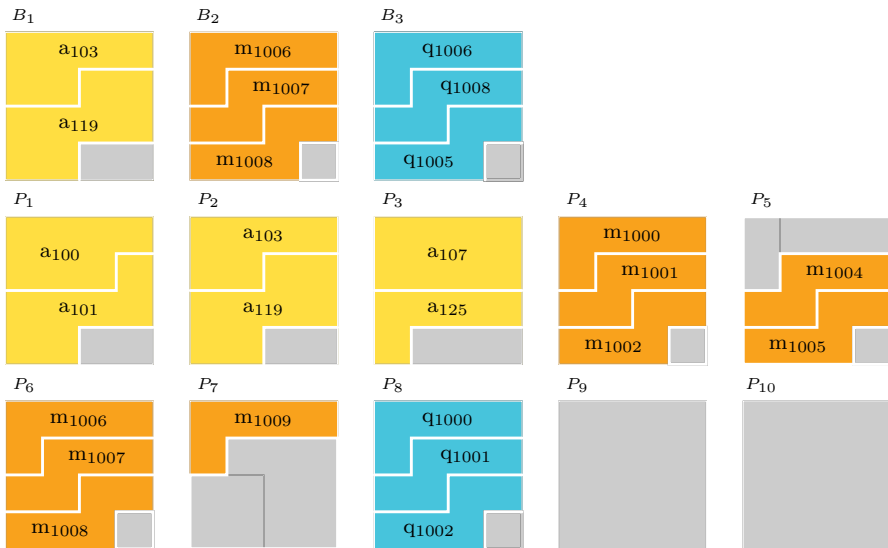
Block Nested Loop Join: Minimum Buffer Space



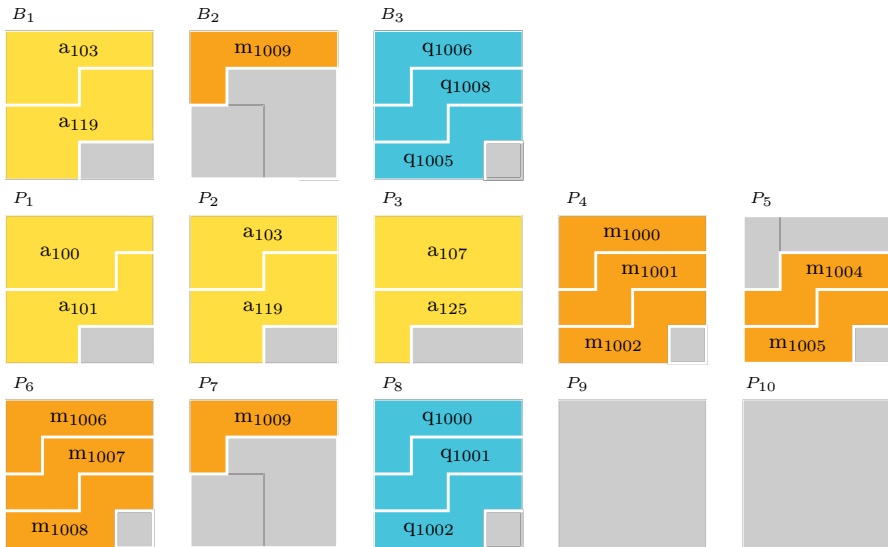
Block Nested Loop Join: Minimum Buffer Space



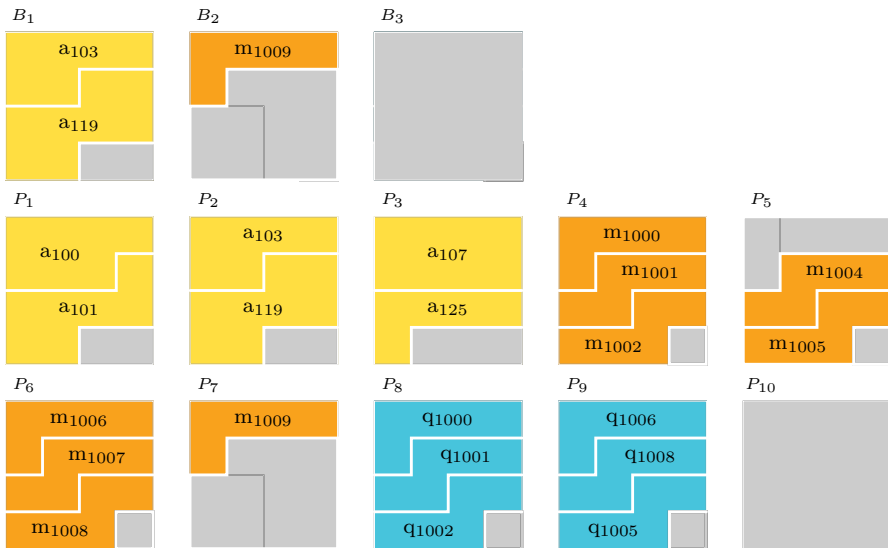
Block Nested Loop Join: Minimum Buffer Space



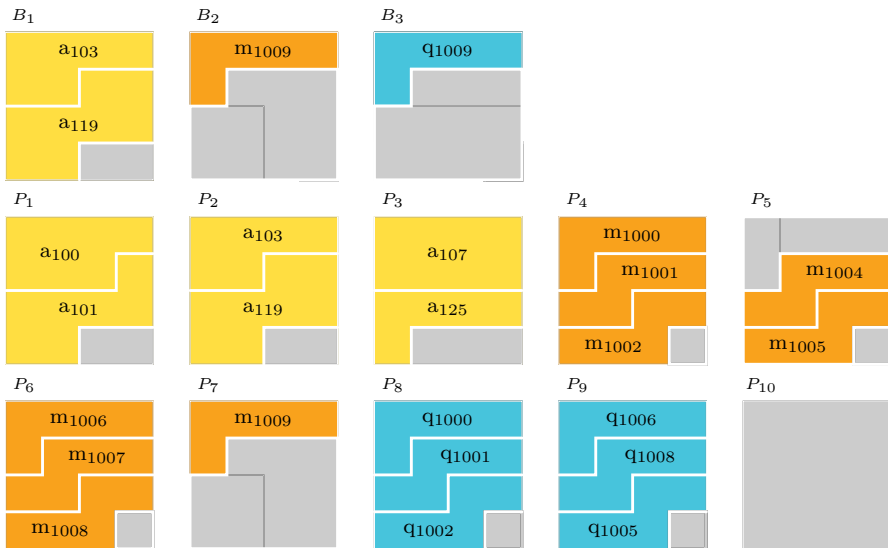
Block Nested Loop Join: Minimum Buffer Space



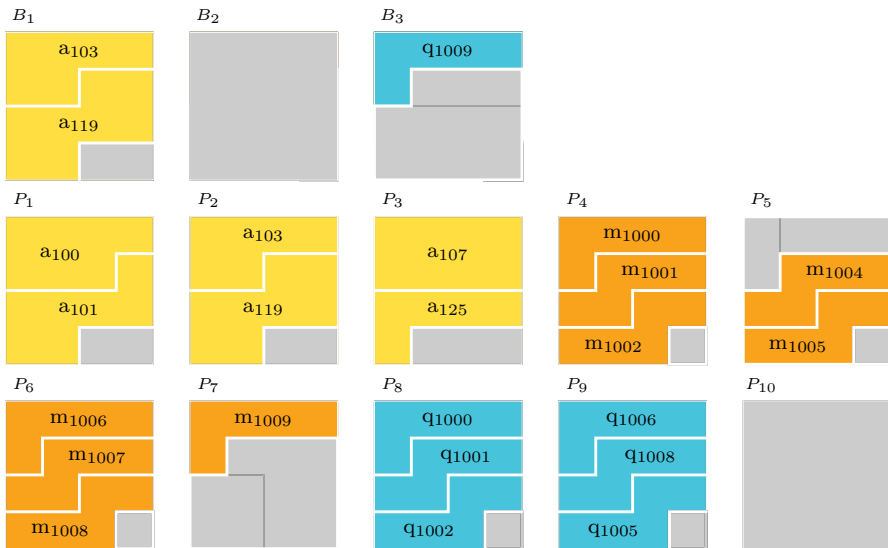
Block Nested Loop Join: Minimum Buffer Space



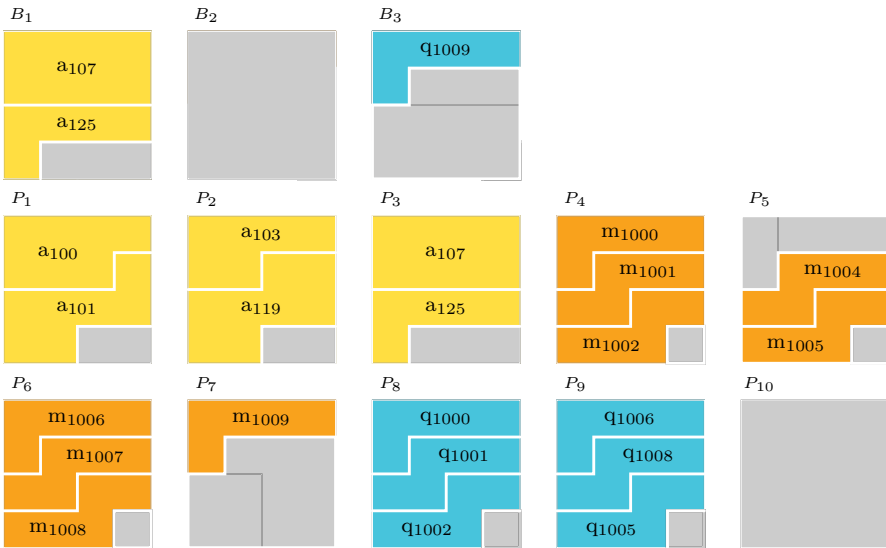
Block Nested Loop Join: Minimum Buffer Space



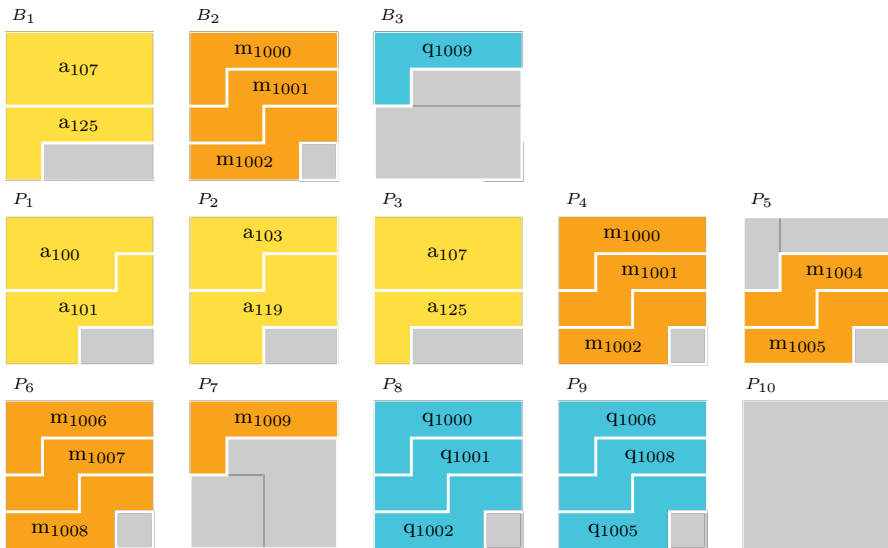
Block Nested Loop Join: Minimum Buffer Space



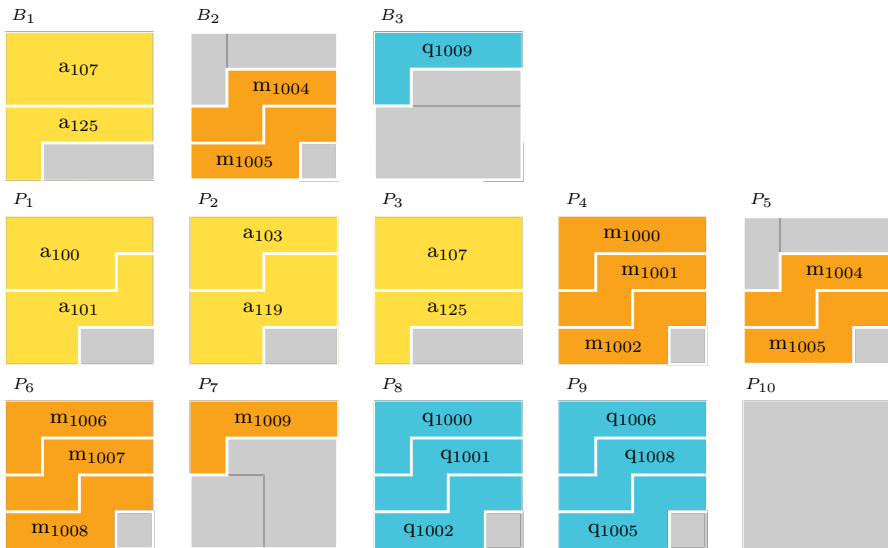
Block Nested Loop Join: Minimum Buffer Space



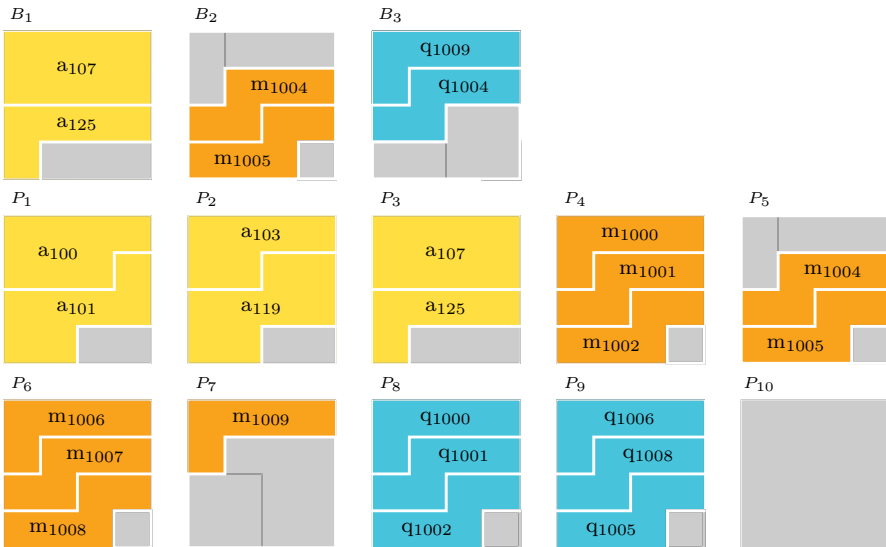
Block Nested Loop Join: Minimum Buffer Space



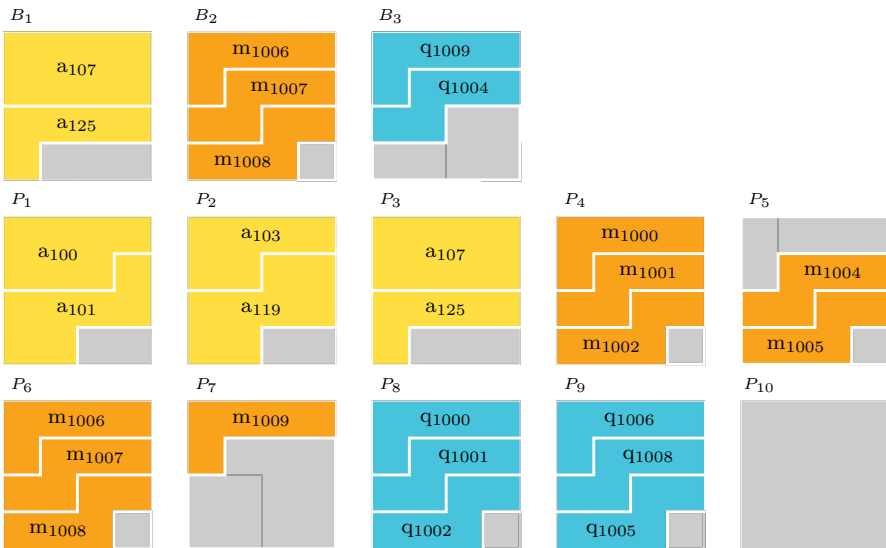
Block Nested Loop Join: Minimum Buffer Space



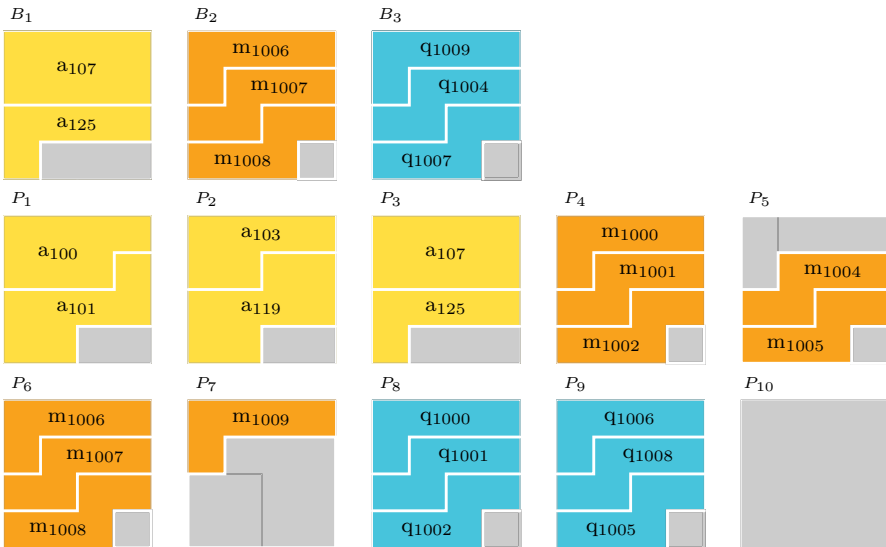
Block Nested Loop Join: Minimum Buffer Space



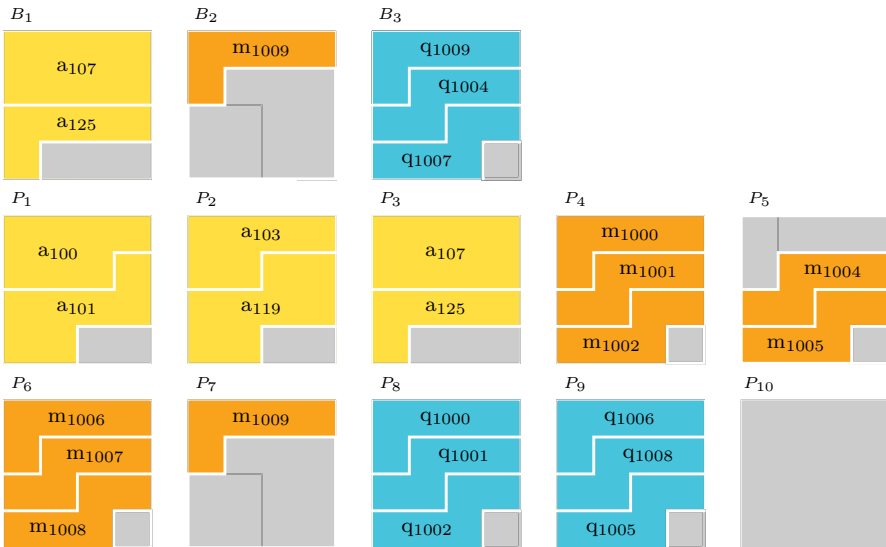
Block Nested Loop Join: Minimum Buffer Space



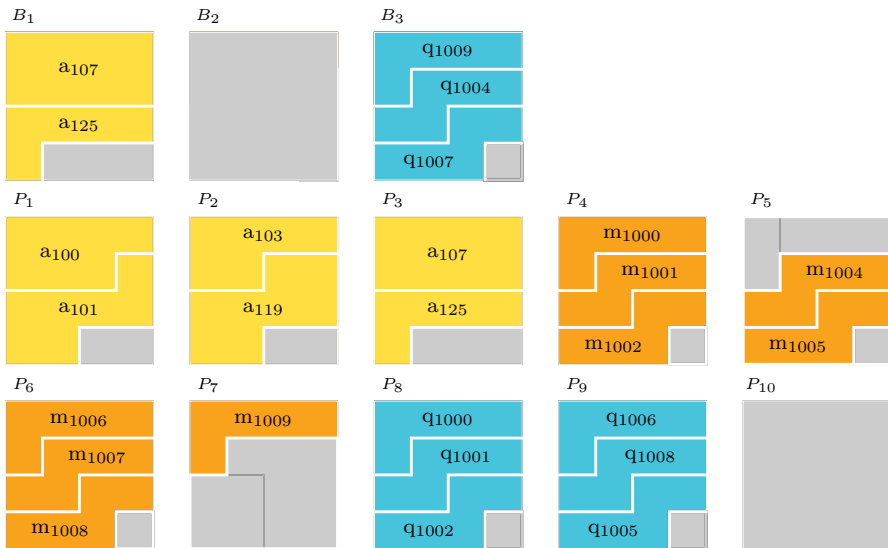
Block Nested Loop Join: Minimum Buffer Space



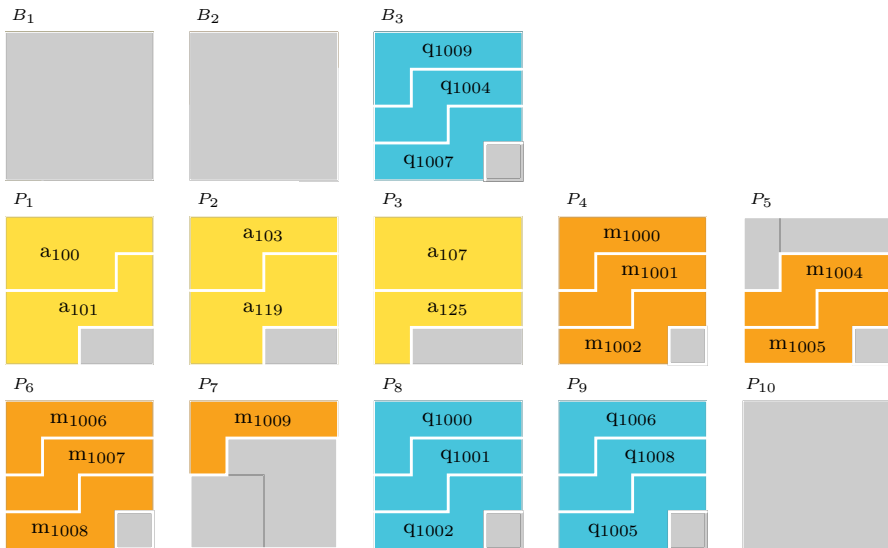
Block Nested Loop Join: Minimum Buffer Space



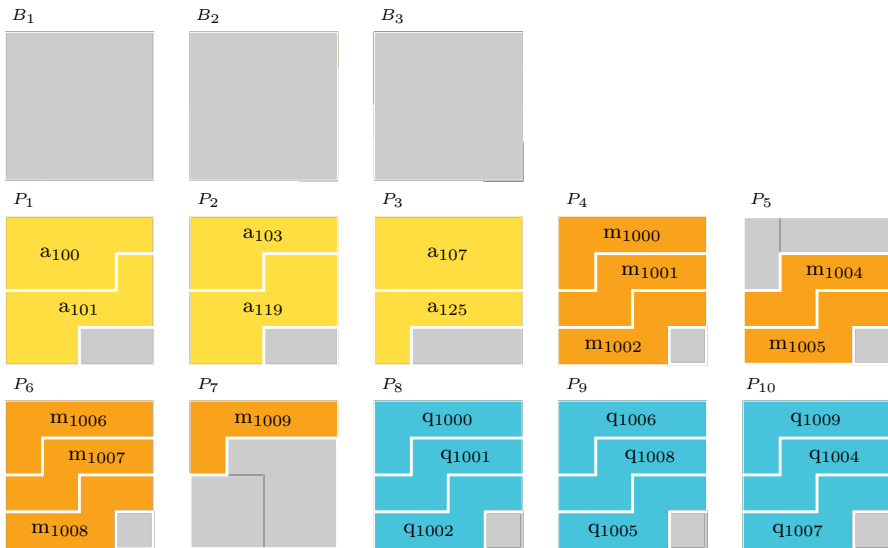
Block Nested Loop Join: Minimum Buffer Space



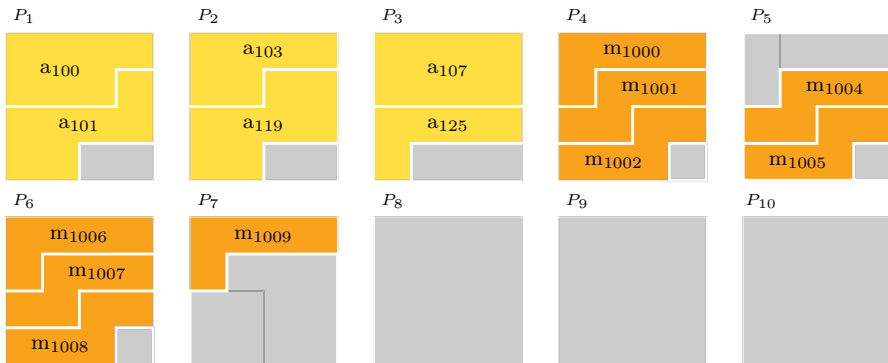
Block Nested Loop Join: Minimum Buffer Space



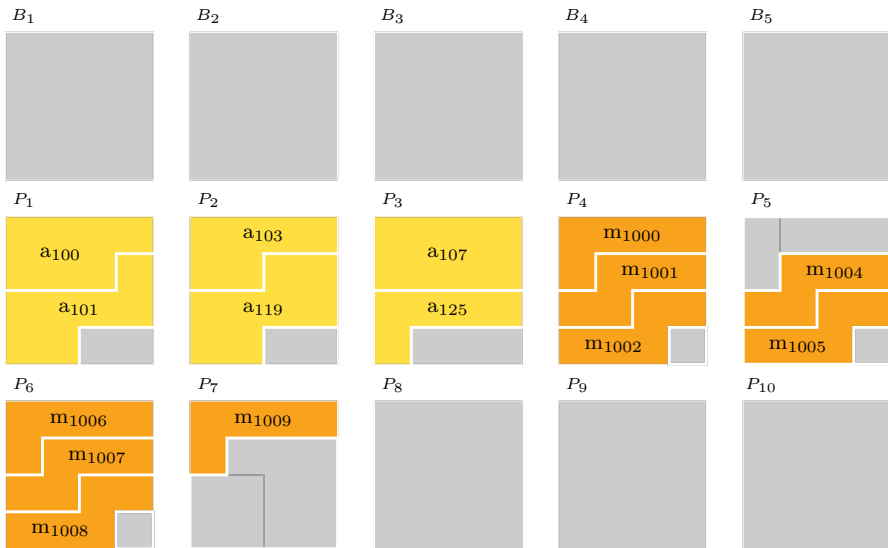
Block Nested Loop Join: Minimum Buffer Space



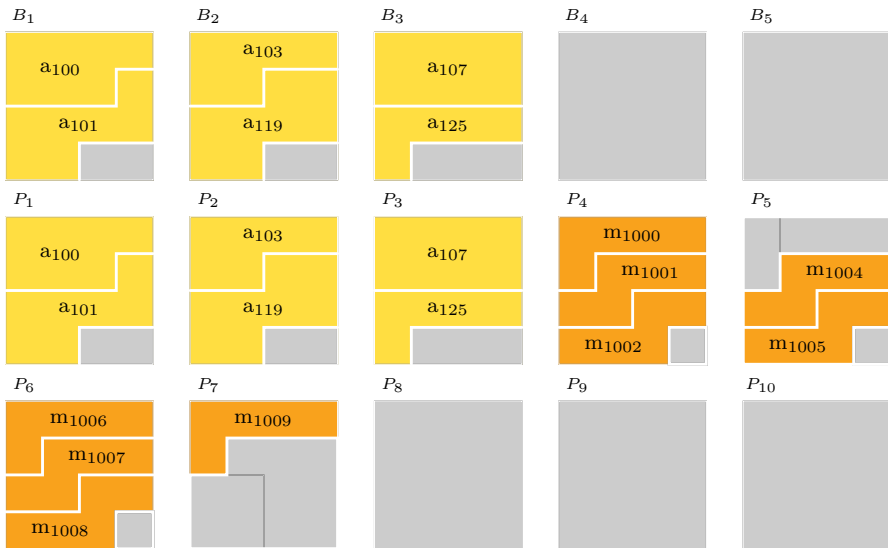
Block Nested Loop Join: Buffer Space for Smallest Relation



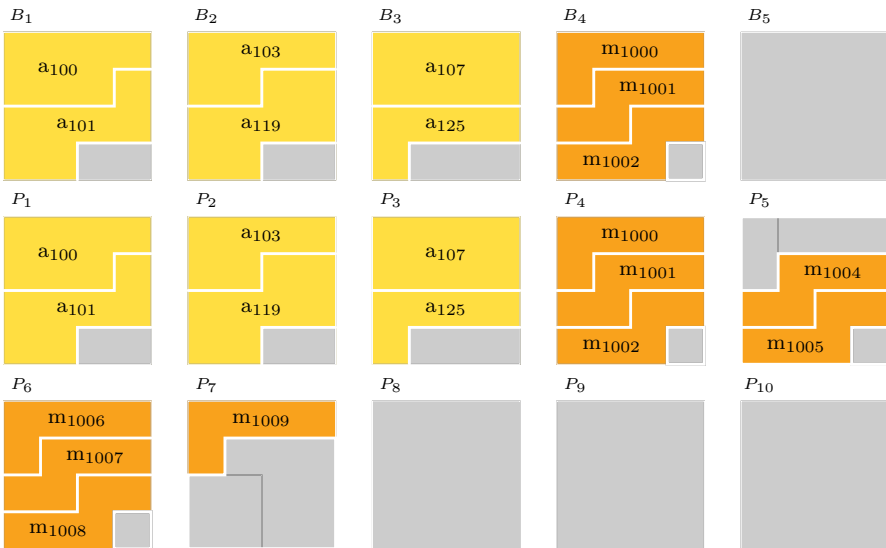
Block Nested Loop Join: Buffer Space for Smallest Relation



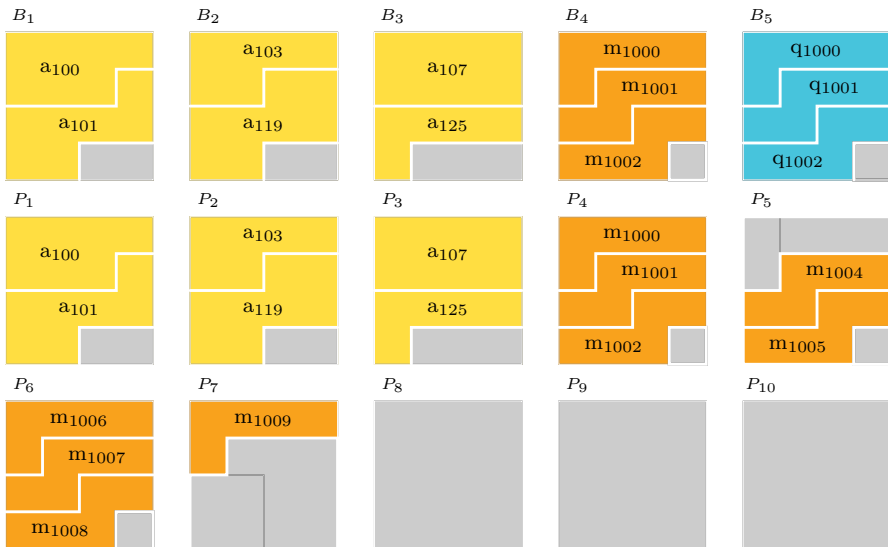
Block Nested Loop Join: Buffer Space for Smallest Relation



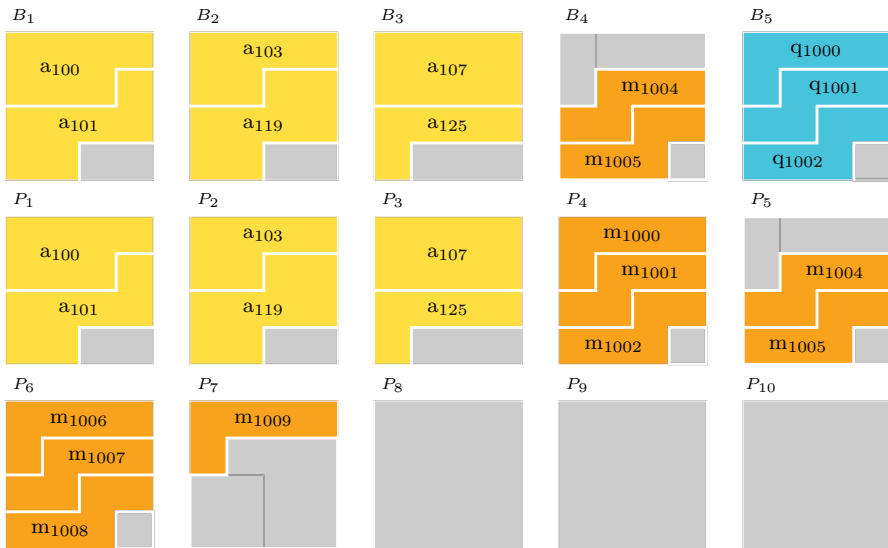
Block Nested Loop Join: Buffer Space for Smallest Relation



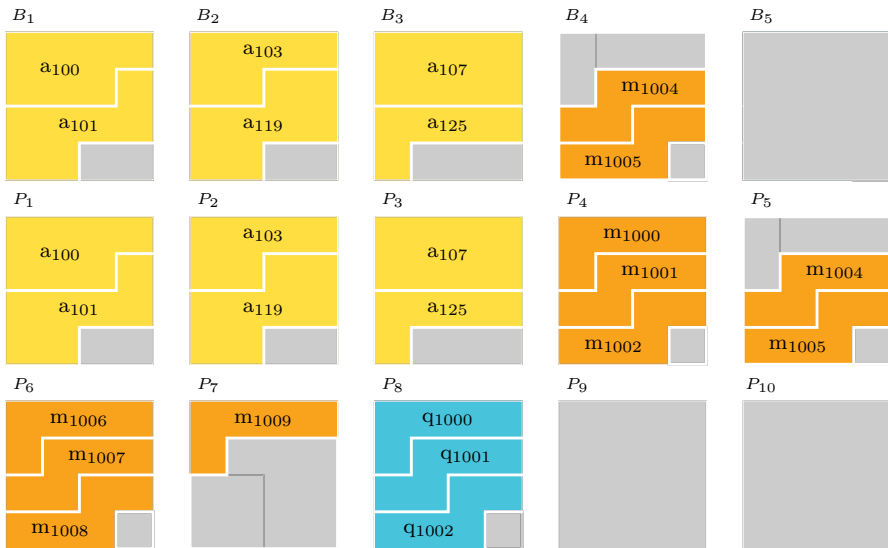
Block Nested Loop Join: Buffer Space for Smallest Relation



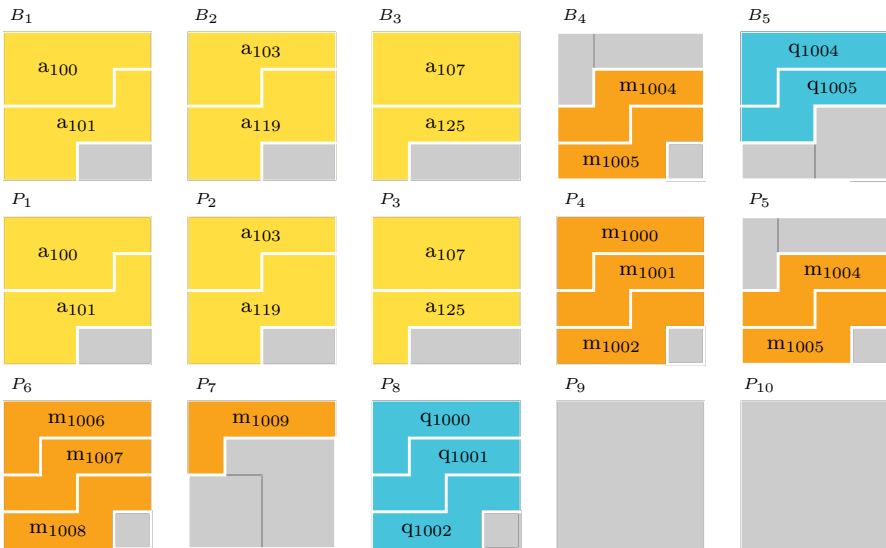
Block Nested Loop Join: Buffer Space for Smallest Relation



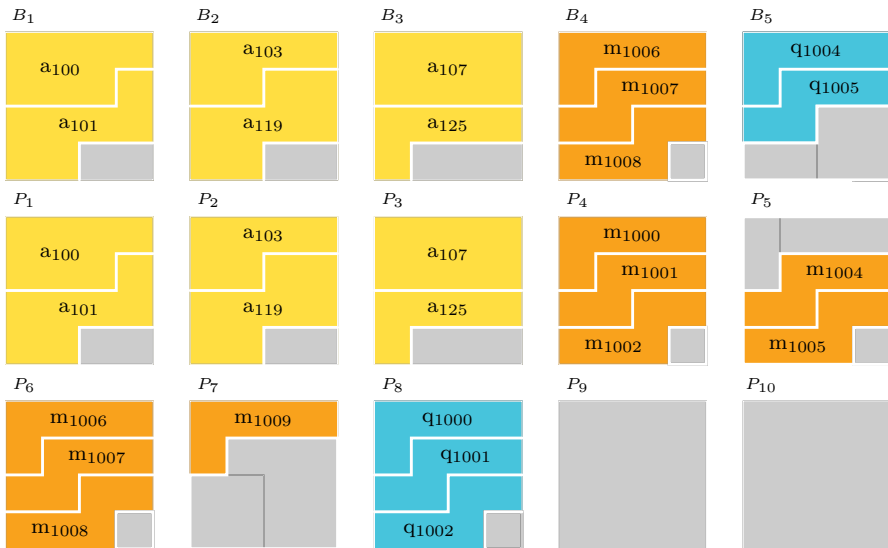
Block Nested Loop Join: Buffer Space for Smallest Relation



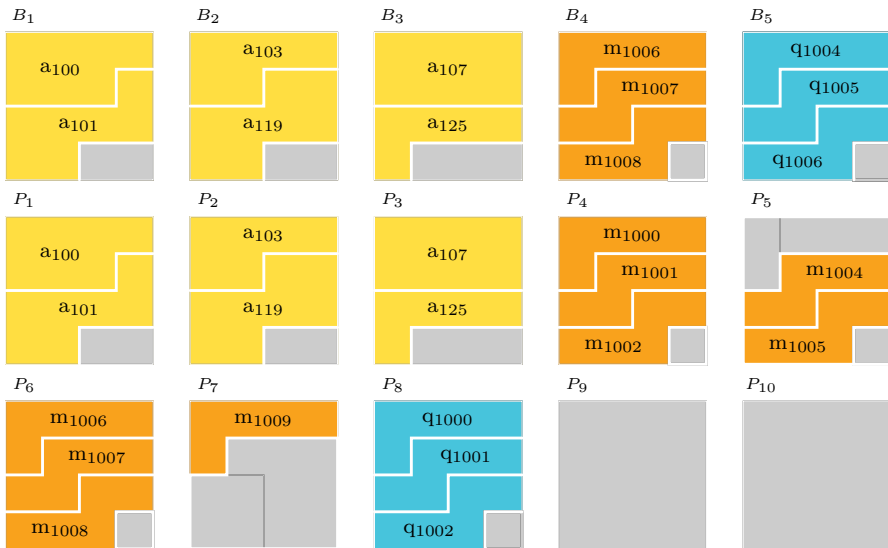
Block Nested Loop Join: Buffer Space for Smallest Relation



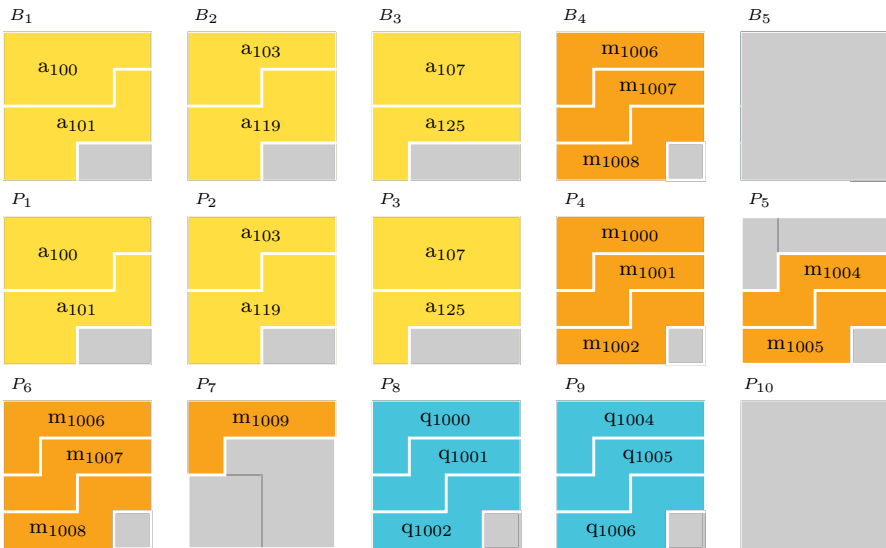
Block Nested Loop Join: Buffer Space for Smallest Relation



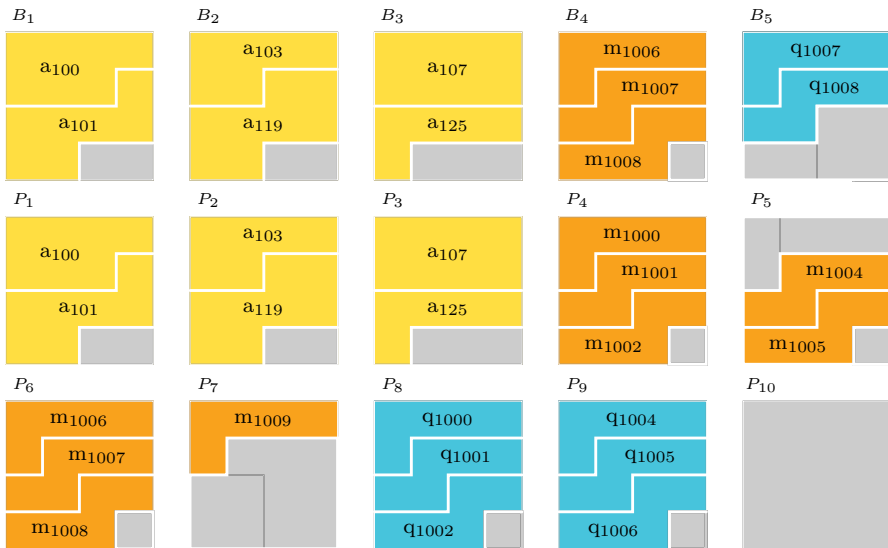
Block Nested Loop Join: Buffer Space for Smallest Relation



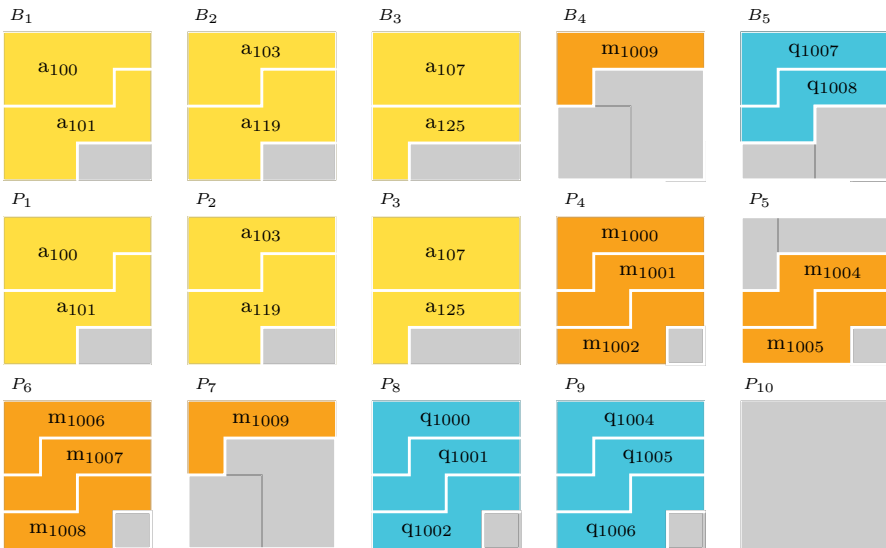
Block Nested Loop Join: Buffer Space for Smallest Relation



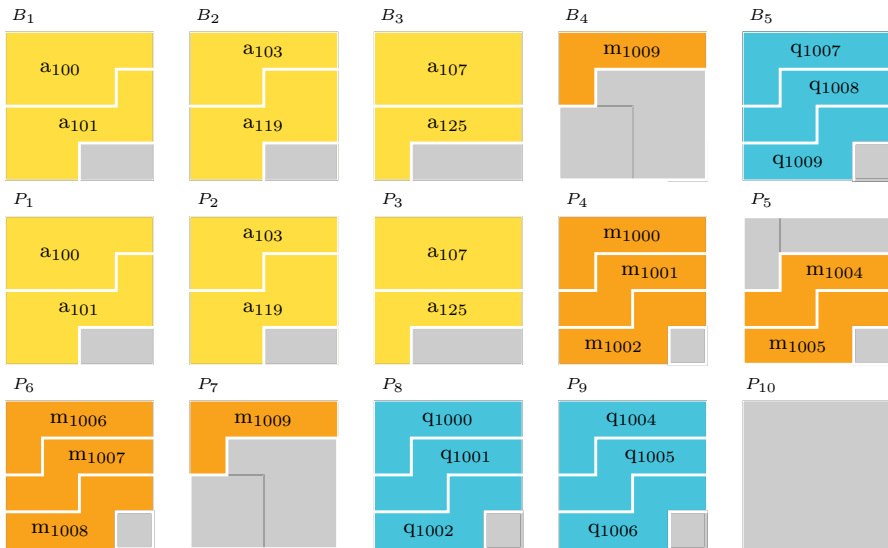
Block Nested Loop Join: Buffer Space for Smallest Relation



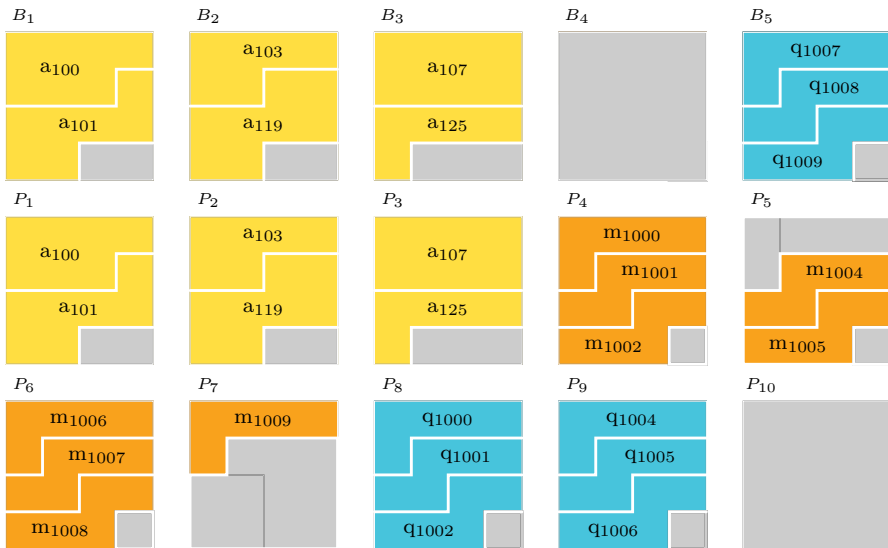
Block Nested Loop Join: Buffer Space for Smallest Relation



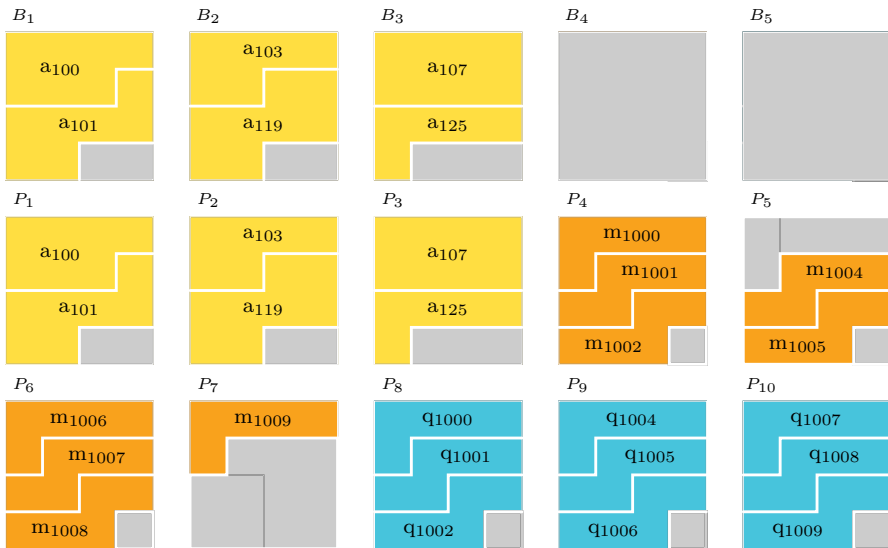
Block Nested Loop Join: Buffer Space for Smallest Relation



Block Nested Loop Join: Buffer Space for Smallest Relation



Block Nested Loop Join: Buffer Space for Smallest Relation



Block Nested Loop Join: Analysis

$$R \stackrel{BNL}{\bowtie} S$$

- If $|B| = 3$ then for each page in R , must read all the pages in S
Number of disc page reads is

$$|Pages(R)| + |Pages(R)| \cdot |Pages(S)|$$

Block Nested Loop Join: Analysis

$$R \stackrel{BNL}{\bowtie} S$$

- If $|B| = 3$ then for each page in R , must read all the pages in S
Number of disc page reads is

$$|Pages(R)| + |Pages(R)| \cdot |Pages(S)|$$

- If $|B| = |Pages(R)| + 2$ then read in all pages of R , and then read in each page of S once to join with all pages of R
Number of disc page reads is

$$|Pages(R)| + |Pages(S)|$$

Block Nested Loop Join: Analysis

$$R \stackrel{BNL}{\bowtie} S$$

- If $|B| = 3$ then for each page in R , must read all the pages in S
Number of disc page reads is

$$|Pages(R)| + |Pages(R)| \cdot |Pages(S)|$$

- If $|B| = |Pages(R)| + 2$ then read in all pages of R , and then read in each page of S once to join with all pages of R
Number of disc page reads is

$$|Pages(R)| + |Pages(S)|$$

- In general, number of page reads is

$$|Pages(R)| + \left\lceil \frac{|Pages(R)|}{|B| - 2} \right\rceil \cdot |Pages(S)|$$

Quiz 2: Performance of BNLJ

If we compare $R \bowtie^{BNL} S$ with $S \bowtie^{BNL} R$, where $|Pages(R)| < |Pages(S)|$

Which will have the lower $NDPR$?

A

$R \bowtie^{BNL} S$

B

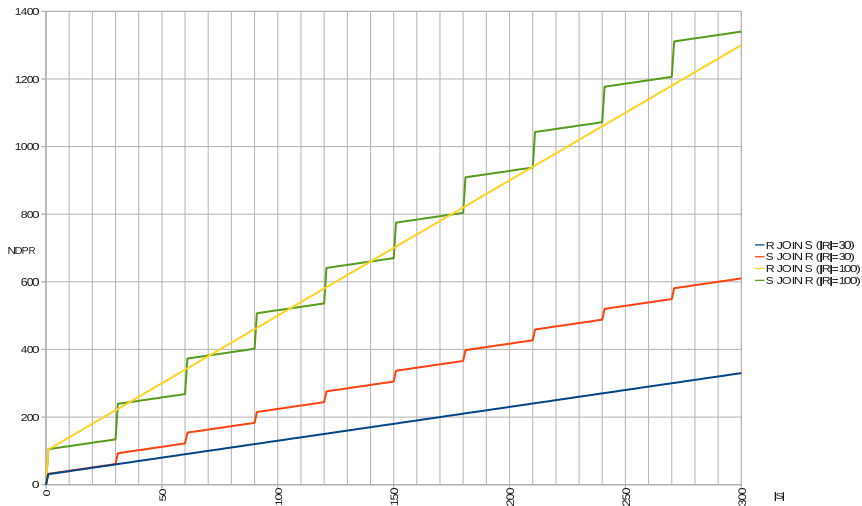
$S \bowtie^{BNL} R$

C

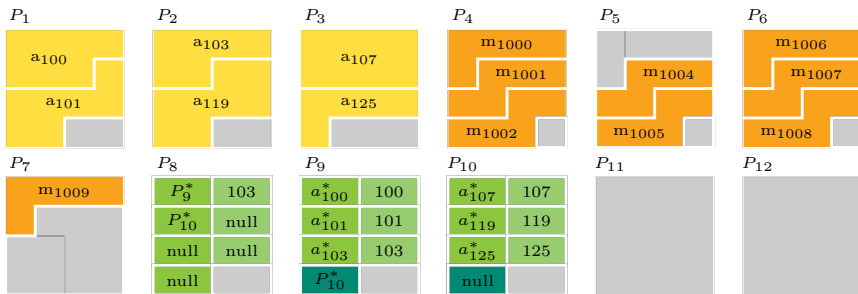
Neither, the two are equal

D

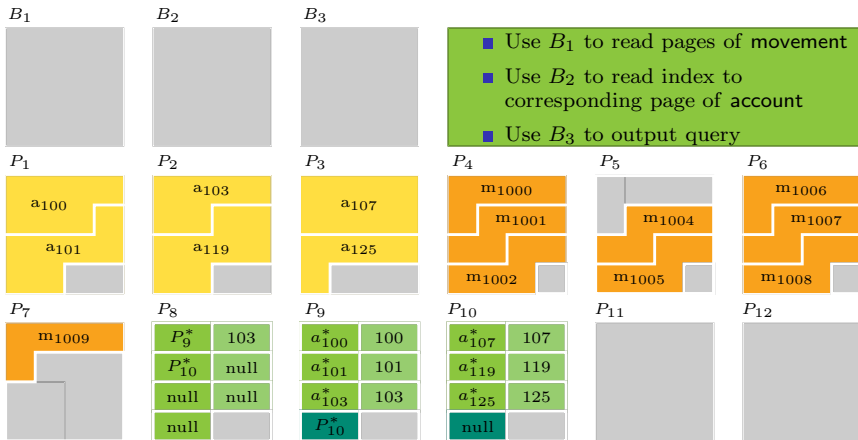
Sometimes $R \bowtie^{BNL} S$, sometimes $S \bowtie^{BNL} R$

Performance of BNLJ, with $|B| = 32$ 

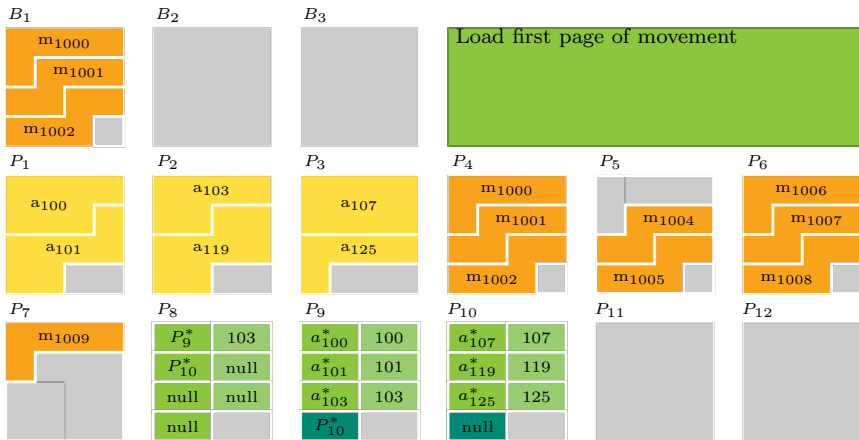
Index Nested Loop Join: Minimum Buffer Space



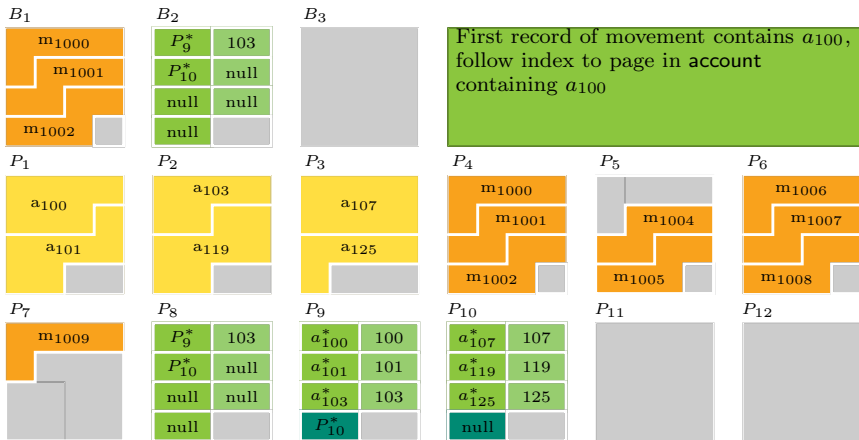
Index Nested Loop Join: Minimum Buffer Space



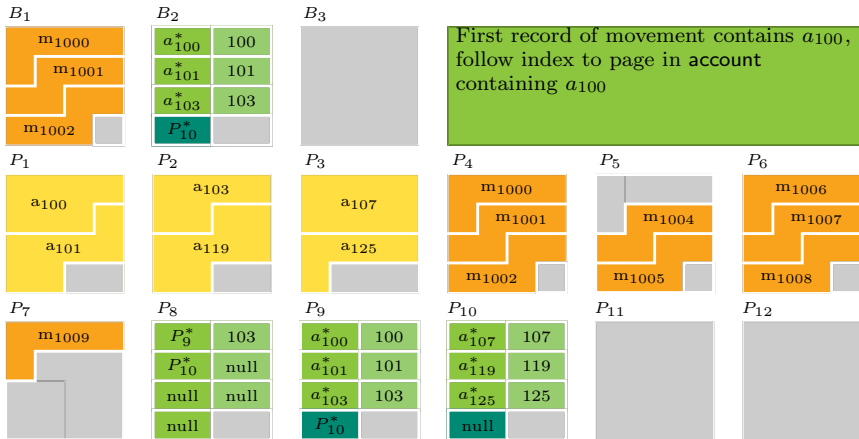
Index Nested Loop Join: Minimum Buffer Space



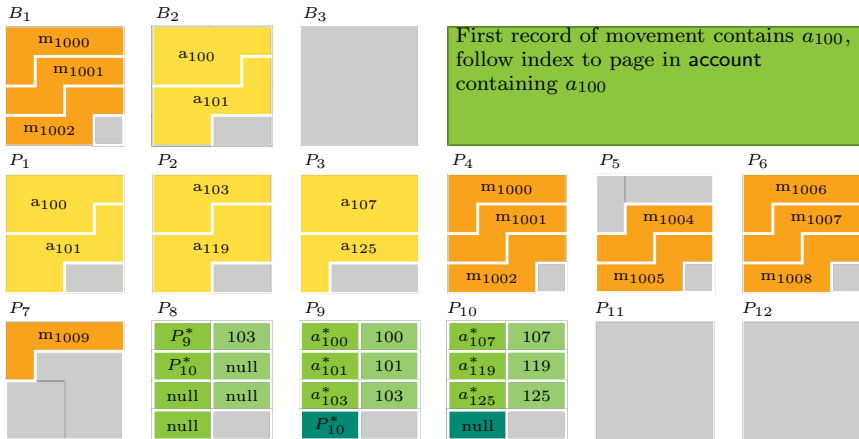
Index Nested Loop Join: Minimum Buffer Space



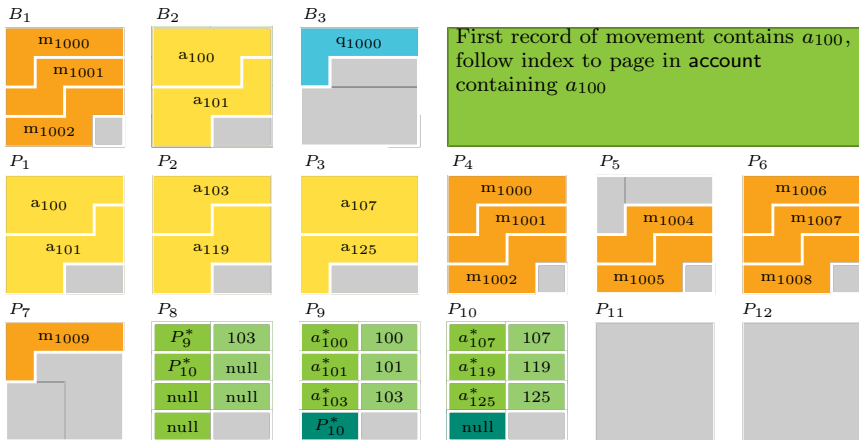
Index Nested Loop Join: Minimum Buffer Space



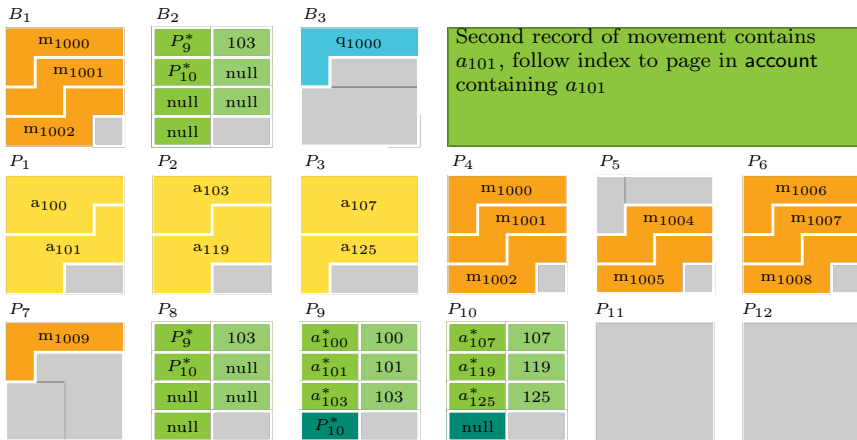
Index Nested Loop Join: Minimum Buffer Space



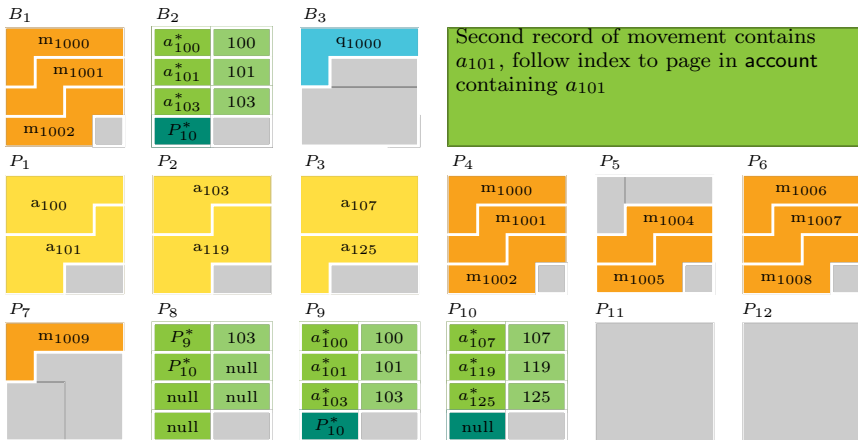
Index Nested Loop Join: Minimum Buffer Space



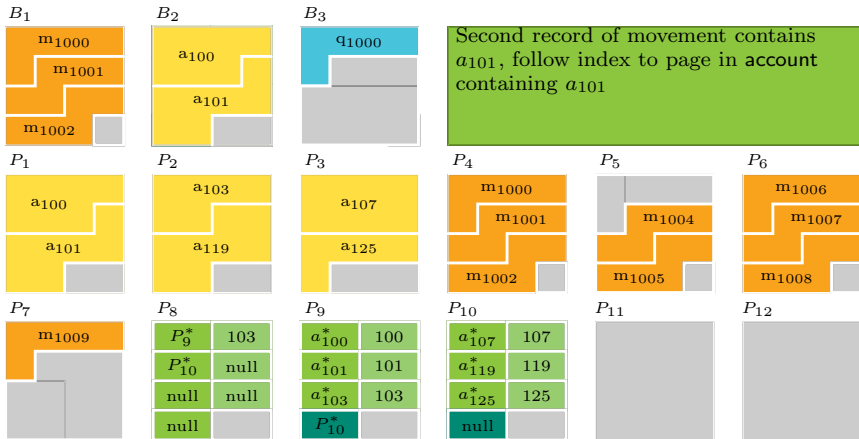
Index Nested Loop Join: Minimum Buffer Space



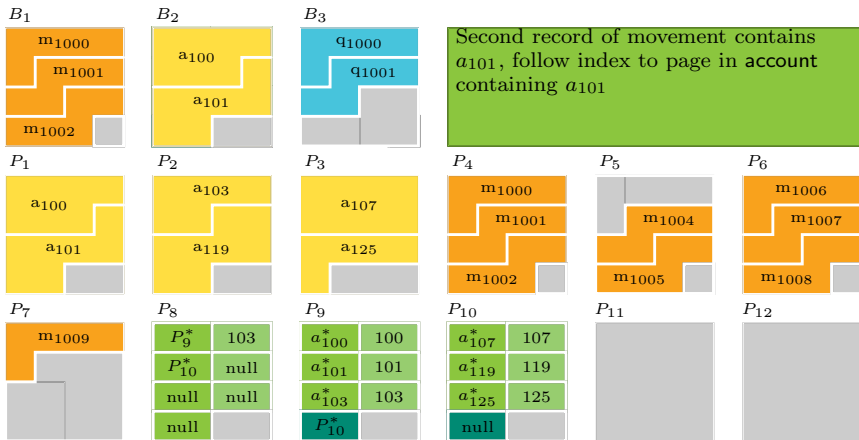
Index Nested Loop Join: Minimum Buffer Space



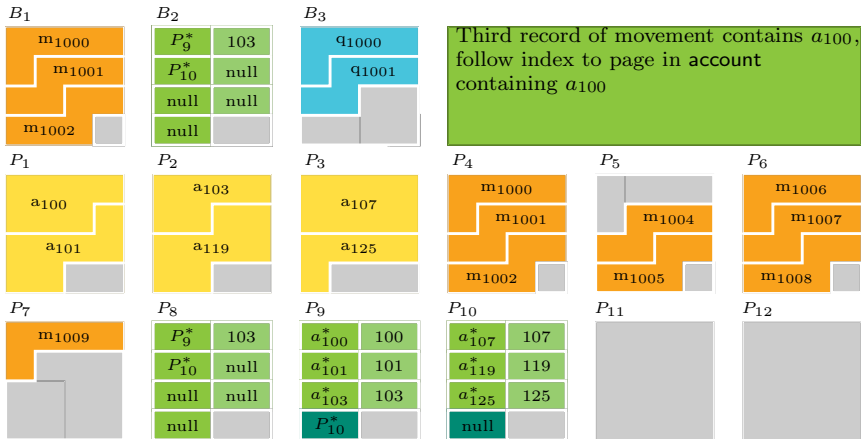
Index Nested Loop Join: Minimum Buffer Space



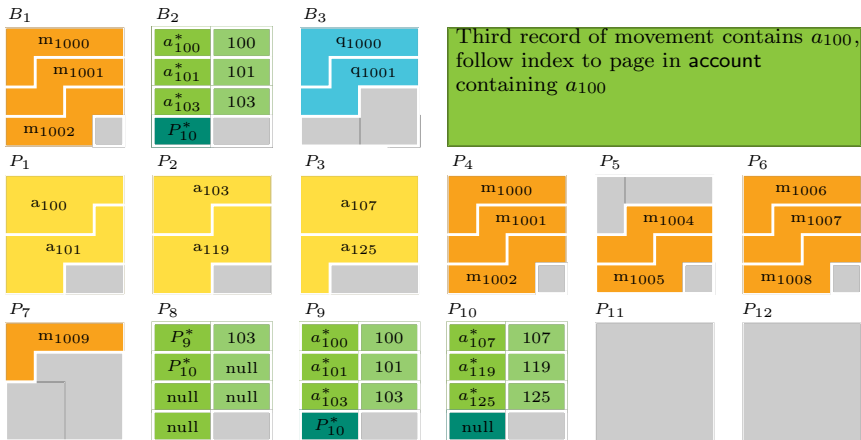
Index Nested Loop Join: Minimum Buffer Space



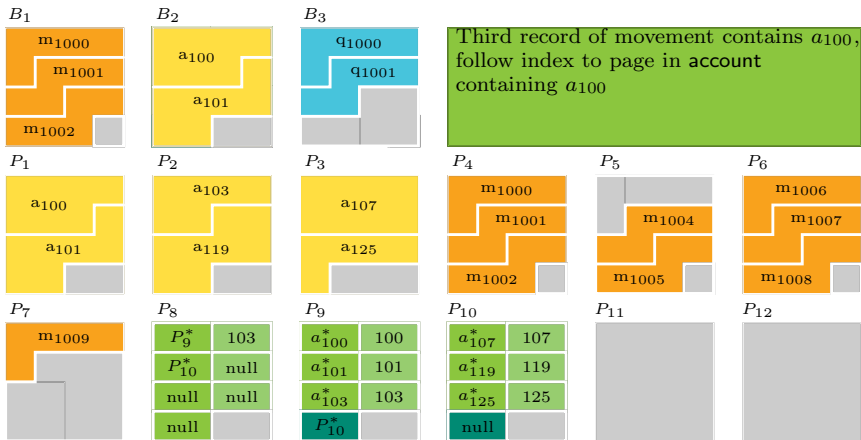
Index Nested Loop Join: Minimum Buffer Space



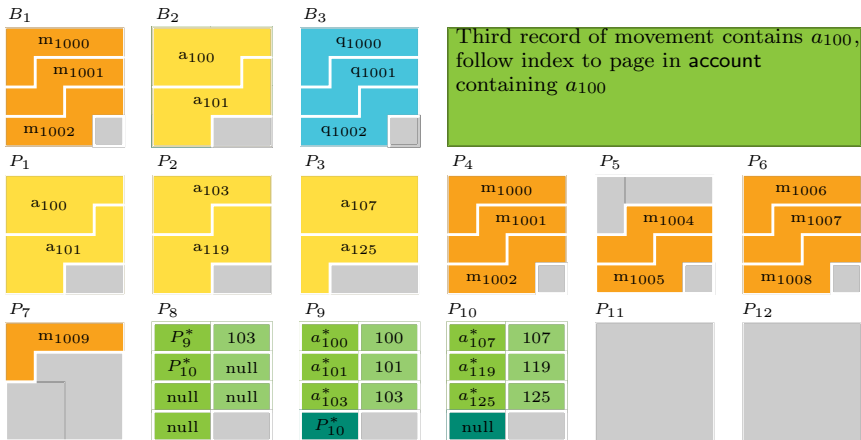
Index Nested Loop Join: Minimum Buffer Space



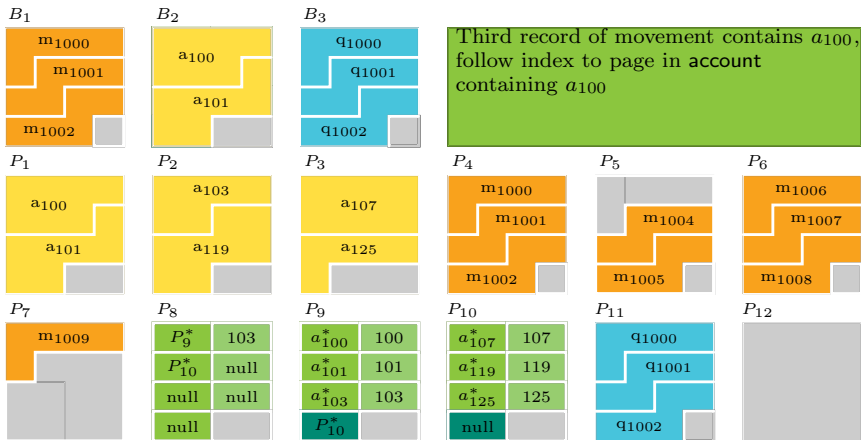
Index Nested Loop Join: Minimum Buffer Space



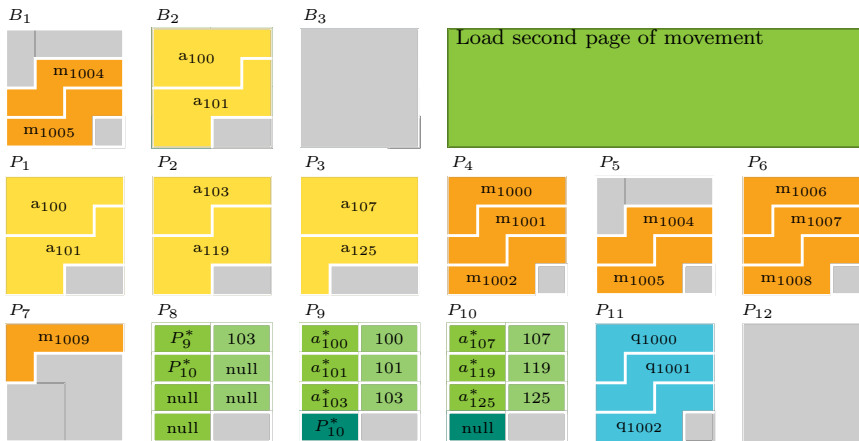
Index Nested Loop Join: Minimum Buffer Space



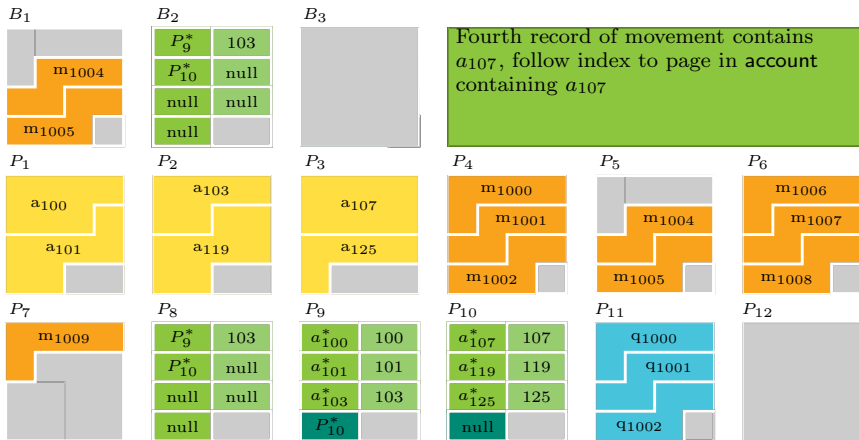
Index Nested Loop Join: Minimum Buffer Space



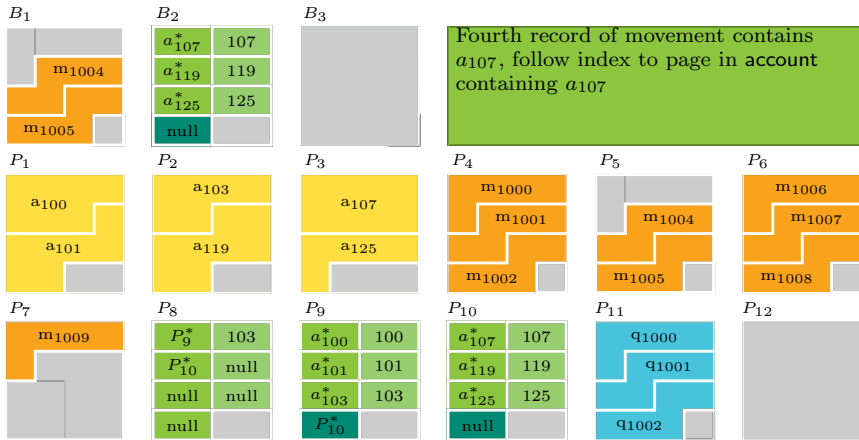
Index Nested Loop Join: Minimum Buffer Space



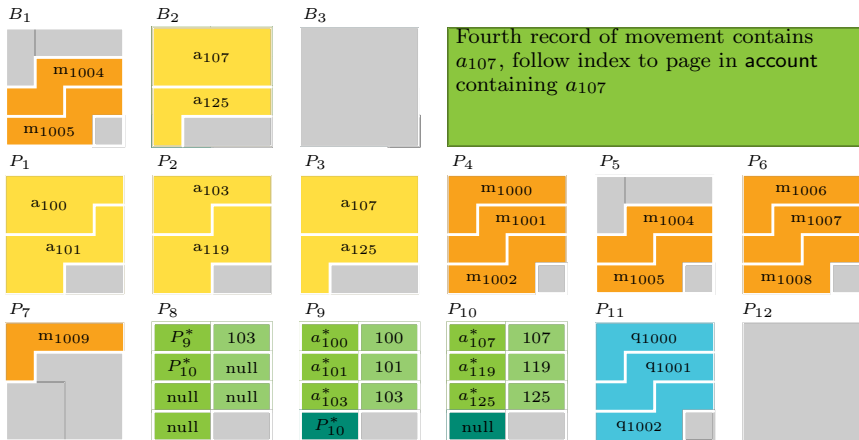
Index Nested Loop Join: Minimum Buffer Space



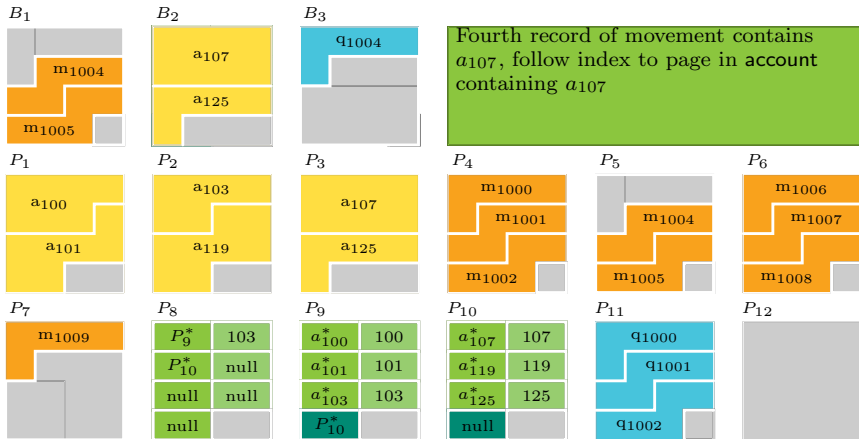
Index Nested Loop Join: Minimum Buffer Space



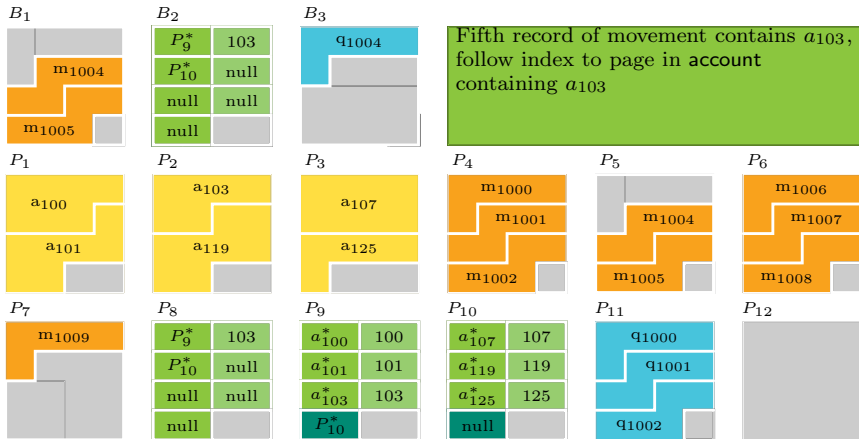
Index Nested Loop Join: Minimum Buffer Space



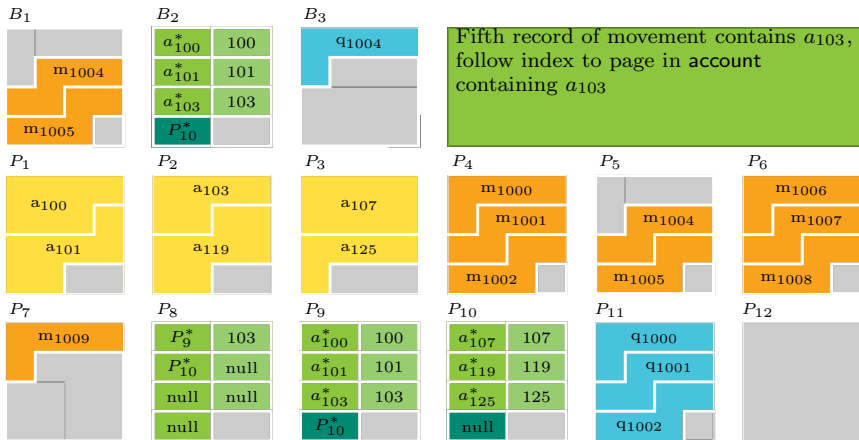
Index Nested Loop Join: Minimum Buffer Space



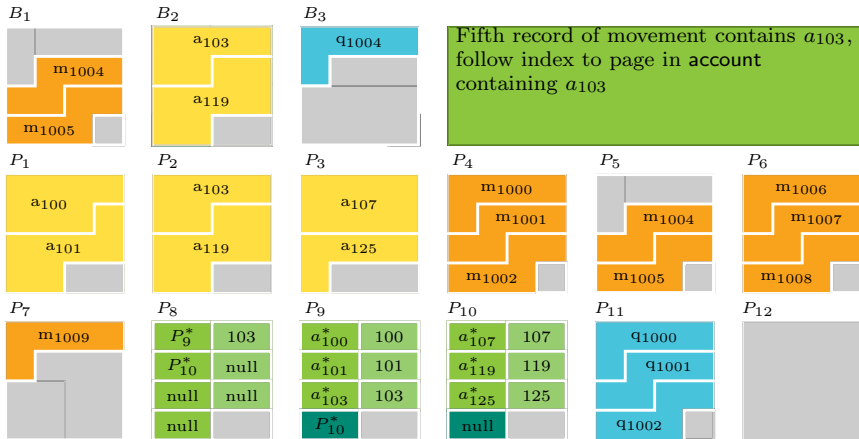
Index Nested Loop Join: Minimum Buffer Space



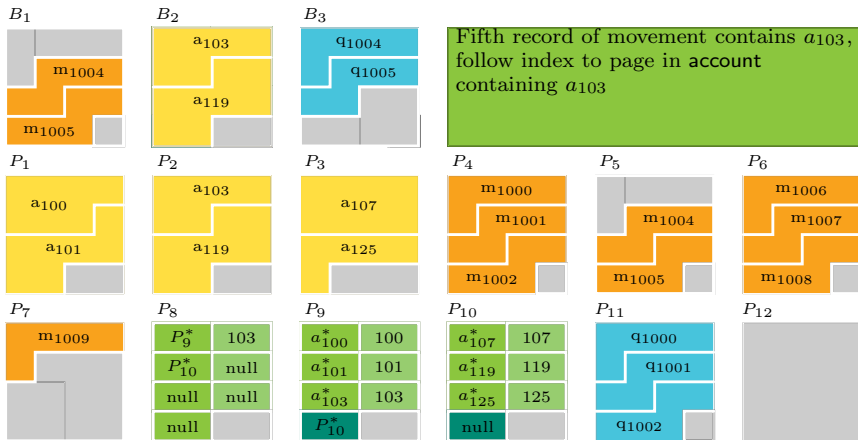
Index Nested Loop Join: Minimum Buffer Space



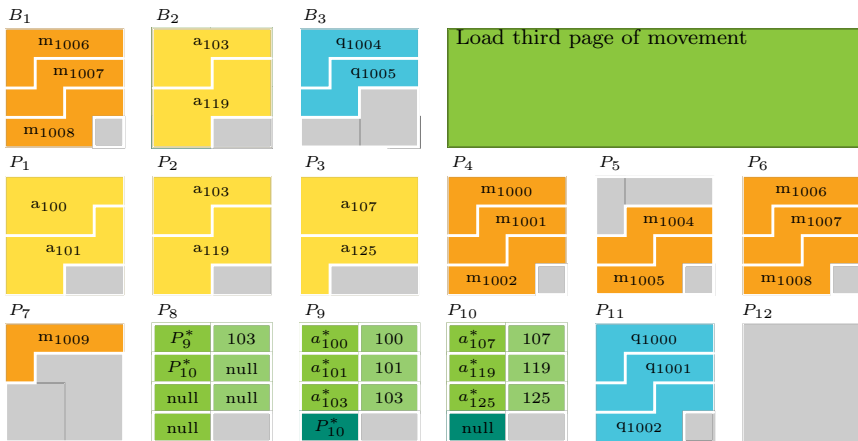
Index Nested Loop Join: Minimum Buffer Space



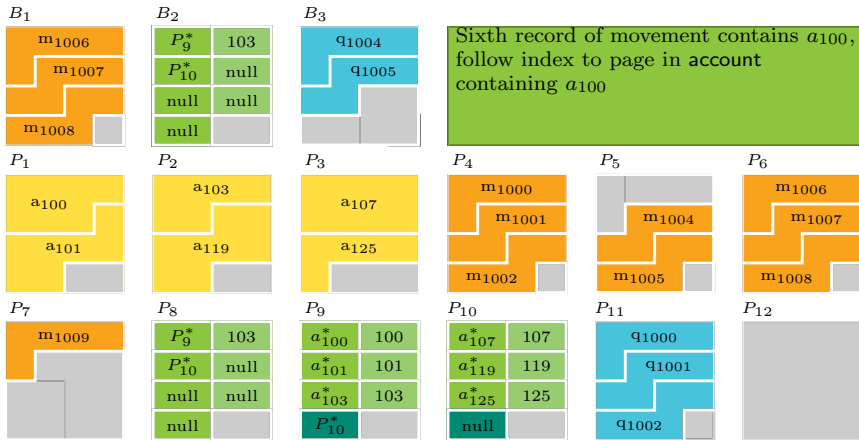
Index Nested Loop Join: Minimum Buffer Space



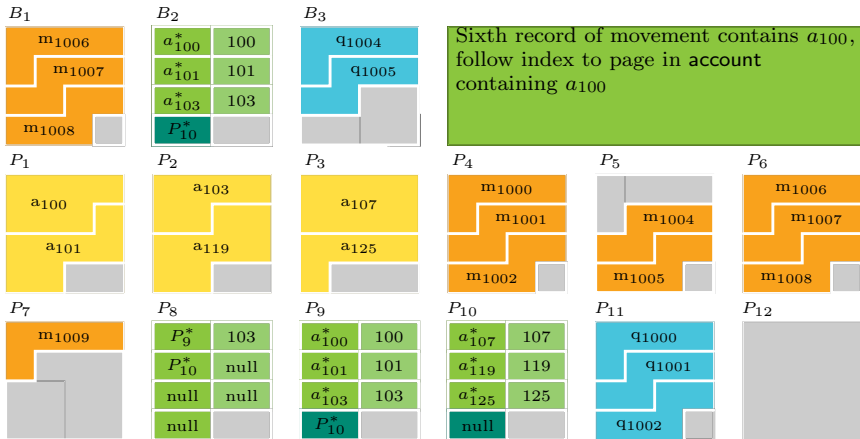
Index Nested Loop Join: Minimum Buffer Space



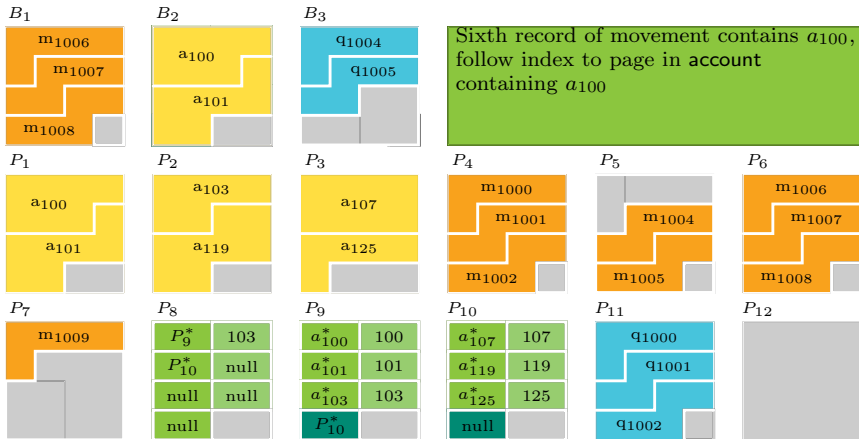
Index Nested Loop Join: Minimum Buffer Space



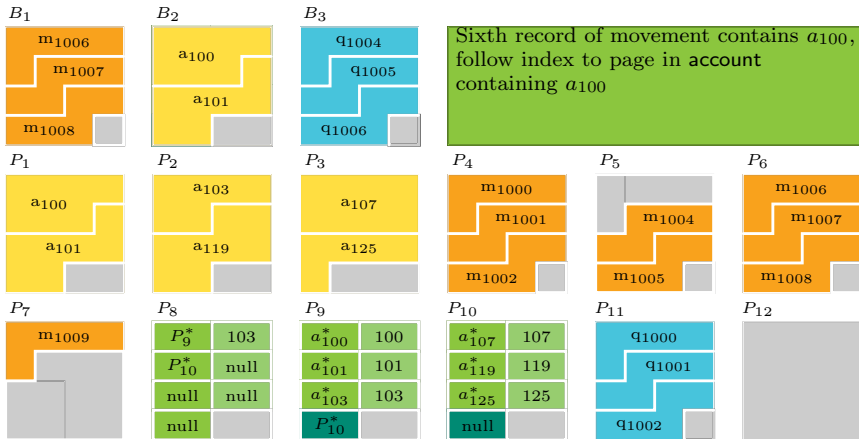
Index Nested Loop Join: Minimum Buffer Space



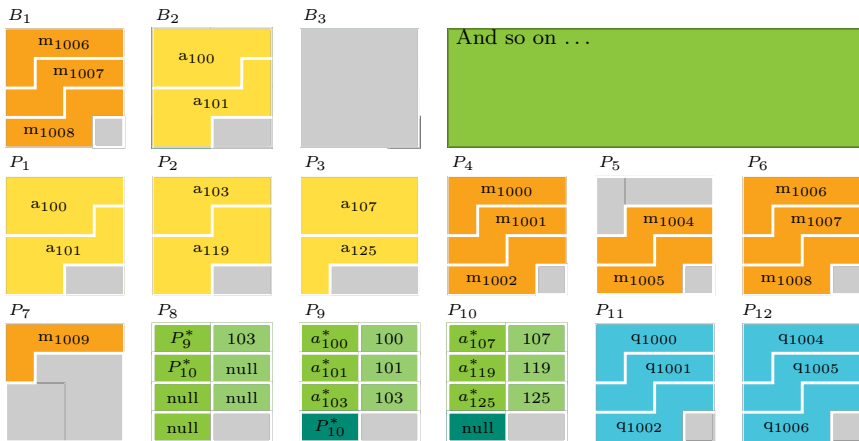
Index Nested Loop Join: Minimum Buffer Space



Index Nested Loop Join: Minimum Buffer Space



Index Nested Loop Join: Minimum Buffer Space



Quiz 3: Choosing a Query Plan

branch		
sortcode	bname	cash
56	'Wimbledon'	94340.45
34	'Goodge St'	8900.67
67	'Strand'	34005.00

account				
no	type	cname	rate	sortcode
100	'current'	'McBrien, P.'	NULL	67
101	'deposit'	'McBrien, P.'	5.25	67
103	'current'	'Boyd, M.'	NULL	34
107	'current'	'Poulovassilis, A.'	NULL	56
119	'deposit'	'Poulovassilis, A.'	5.50	56
125	'current'	'Bailey, J.'	NULL	56

If we execute the following query on the above schema

```
SELECT bname, cname
FROM branch NATURAL JOIN account
```

Which Query Plan is not available?

A

BNL
account \bowtie branch

B

BNL
branch \bowtie account

C

INL
account \bowtie branch

D

INL
branch \bowtie account

Index Nested Loop Join: Analysis

$R \bowtie^{INLJ} S$, and I_S is a B^+ -tree index over the join attribute of S

- If $|B| = 3$

$$NDPR = |Pages(R)| + |R|.Depth(I_S) + |R \bowtie S|$$

Index Nested Loop Join: Analysis

$R \bowtie^{INLJ} S$, and I_S is a B^+ -tree index over the join attribute of S

- If $|B| = 3$

$$NDPR = |Pages(R)| + |R|.Depth(I_S) + |R \bowtie S|$$

- If $|B| = |Pages(I_S)| + 3$

$$NDPR = |Pages(R)| + |Pages(I_S)| + |R \bowtie S|$$

Worksheet: Nested Loop Joins and Index Loop Joins

account

<u>no</u>	type	cname	rate	sortcode
-----------	------	-------	------	----------

movement

<u>mid</u>	no	amount	tdate
------------	----	--------	-------

movement.no → account.no

column_size

column	bytes
no	4
type	10
cname	20
rate	4
sortcode	6
mid	4
amount	8
tdate	8

Logical Query Optimisation

Push project inside joins

$$\pi_{R.A, S.B, \dots}(R \bowtie^{\theta} S) \equiv ((\pi_{A, \dots} R) \bowtie^{\theta} (\pi_{B, \dots} S))$$

if $R.A, S.B, \dots$ includes attributes used in θ

Reduces memory required to hold R, S pairs

Push select inside joins

$$\sigma_P(R \bowtie^{\theta} S) \equiv ((\sigma_P R) \bowtie^{\theta} S)$$

if P only contains attributes from R and not S

Reduces number of R, S pairs to be considered in join

Quiz 4: Choosing an Efficient Query Plan

branch		
sortcode	bname	cash
56	'Wimbledon'	94340.45
34	'Goodge St'	8900.67
67	'Strand'	34005.00

account					
no	type	cname	rate	sortcode	
100	'current'	'McBrien, P.'	NULL	67	
101	'deposit'	'McBrien, P.'	5.25	67	
103	'current'	'Boyd, M.'	NULL	34	
107	'current'	'Poulovassilis, A.'	NULL	56	
119	'deposit'	'Poulovassilis, A.'	5.50	56	
125	'current'	'Bailey, J.'	NULL	56	

Assume that there are 100 rows in `branch`, and 10000 rows in `account`, and we execute the following query on the above schema

```
SELECT bname, cname
FROM branch NATURAL JOIN account
WHERE account=100
```

Which Query Plan has the lowest NDPR?

A

BNL
account \bowtie branch

B

BNL
branch \bowtie account

C

INL
account \bowtie branch

D

INL
branch \bowtie account

Index Sort-Merge Join: Analysis

$R \bowtie^{ISMJ} S$, I_R and I_S are B⁺-tree indexes over the join attributes of R and S

- If $|B| = 3$

$$NDPR = (Depth(I_R) + 1) \cdot |R| + (Depth(I_S) + 1) \cdot |S|$$

Index Sort-Merge Join: Analysis

$R \bowtie^{ISMJ} S$, I_R and I_S are B⁺-tree indexes over the join attributes of R and S

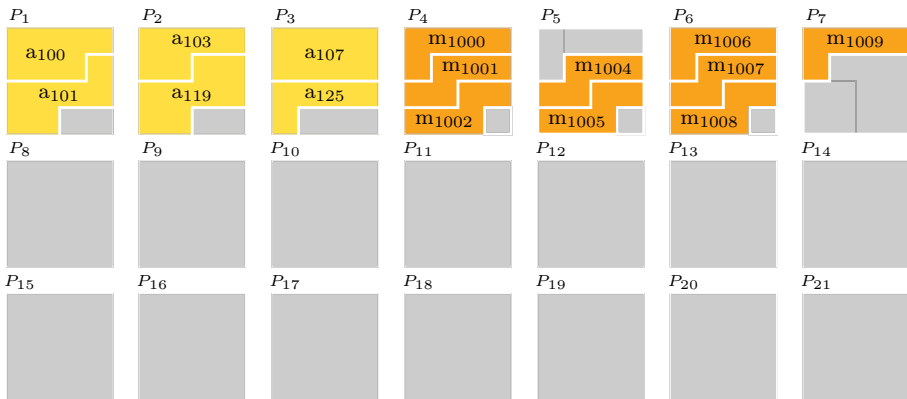
- If $|B| = 3$

$$NDPR = (Depth(I_R) + 1) \cdot |R| + (Depth(I_S) + 1) \cdot |S|$$

- If $|B| = 5$

$$NDPR = |Pages(I_R)| + |R \times S| + |Pages(I_S)| + |S \times R|$$

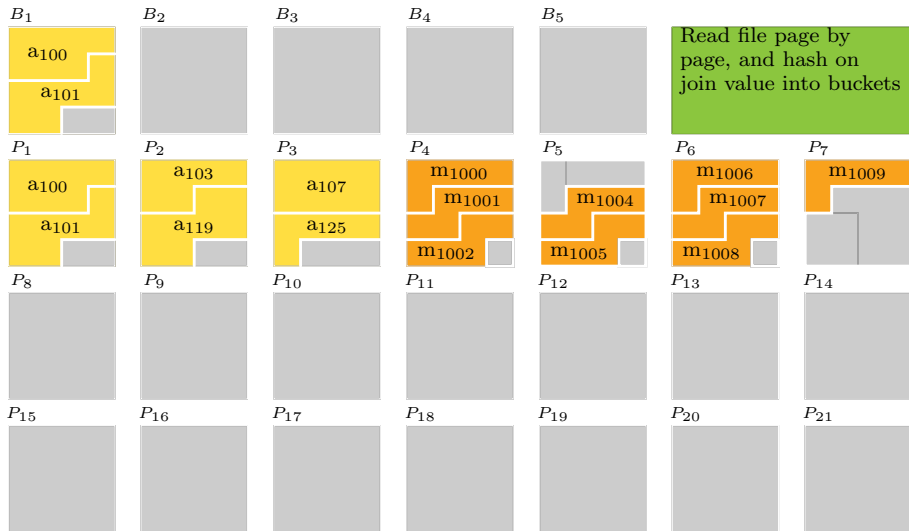
Hash Join



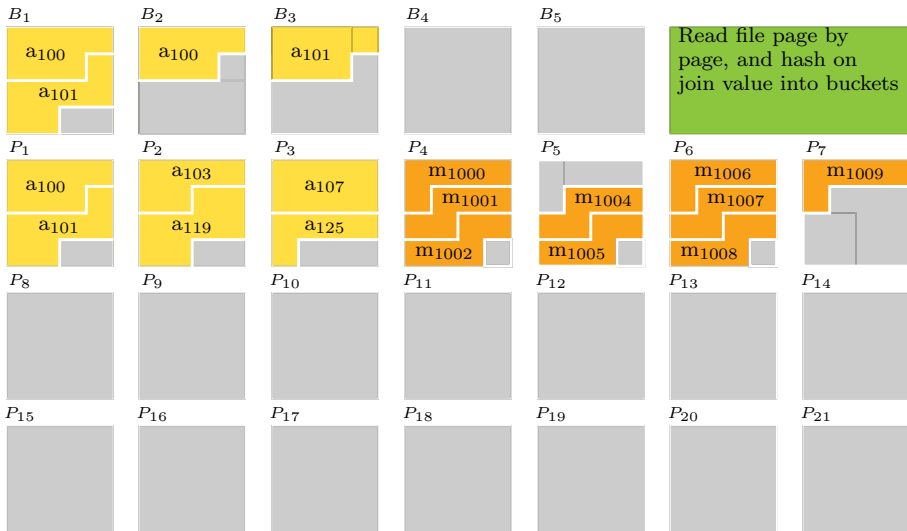
Hash Join



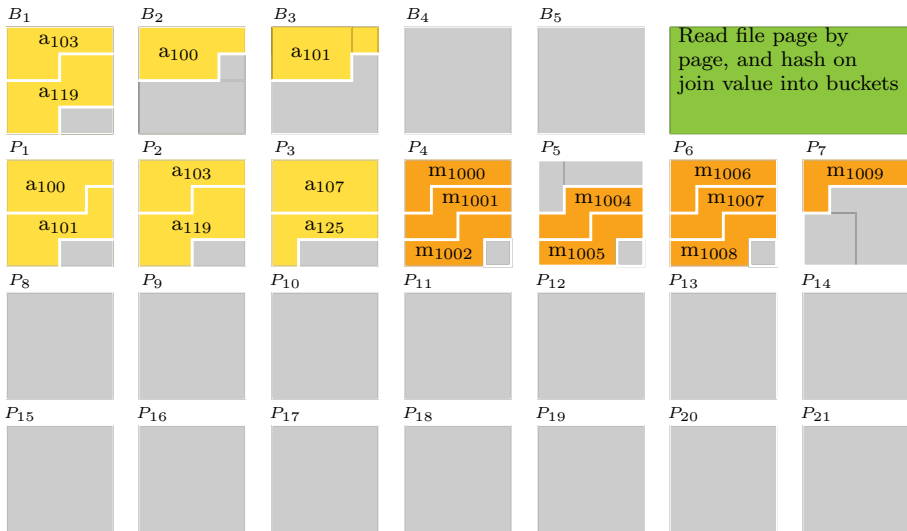
Hash Join



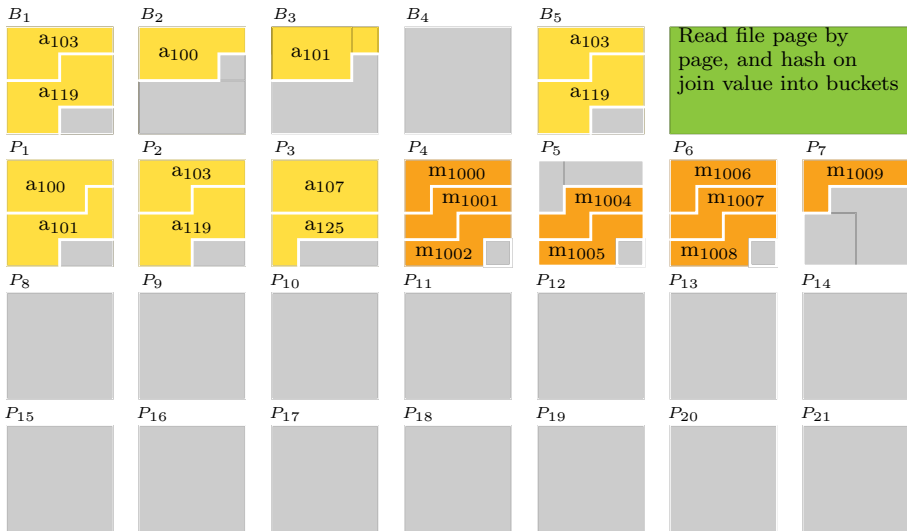
Hash Join



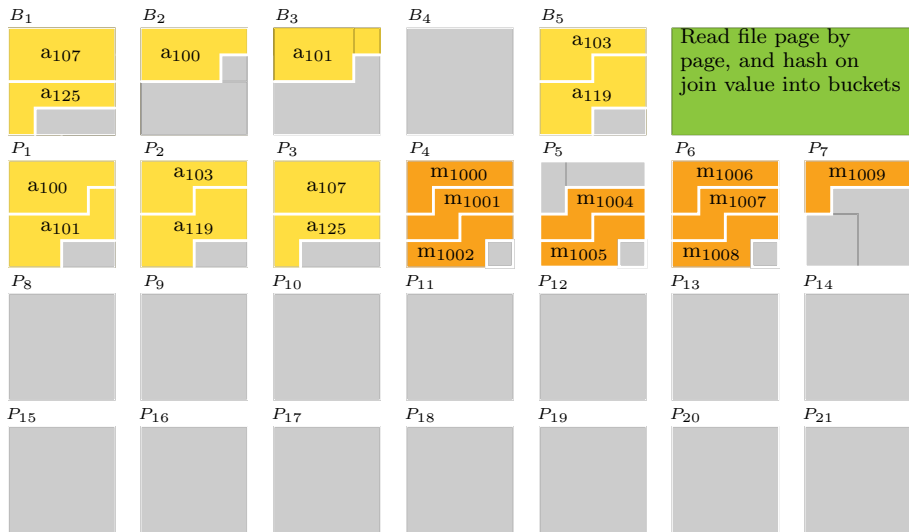
Hash Join



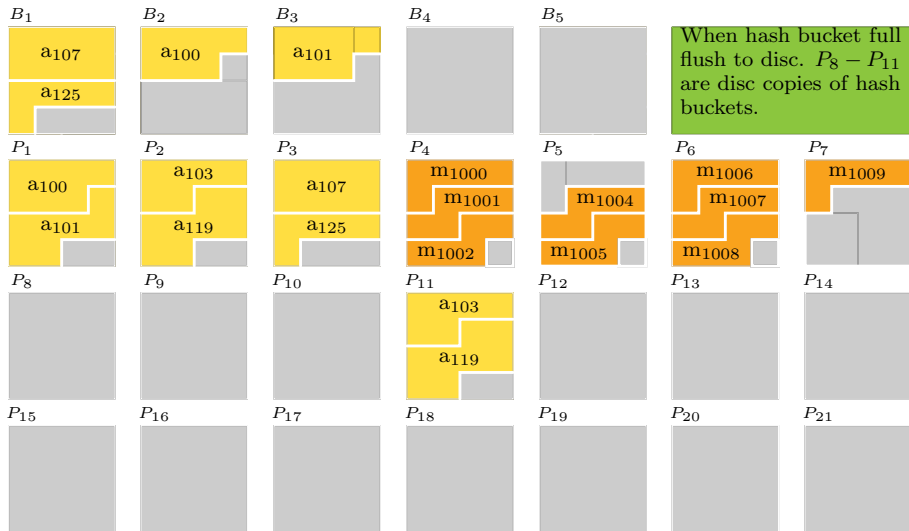
Hash Join



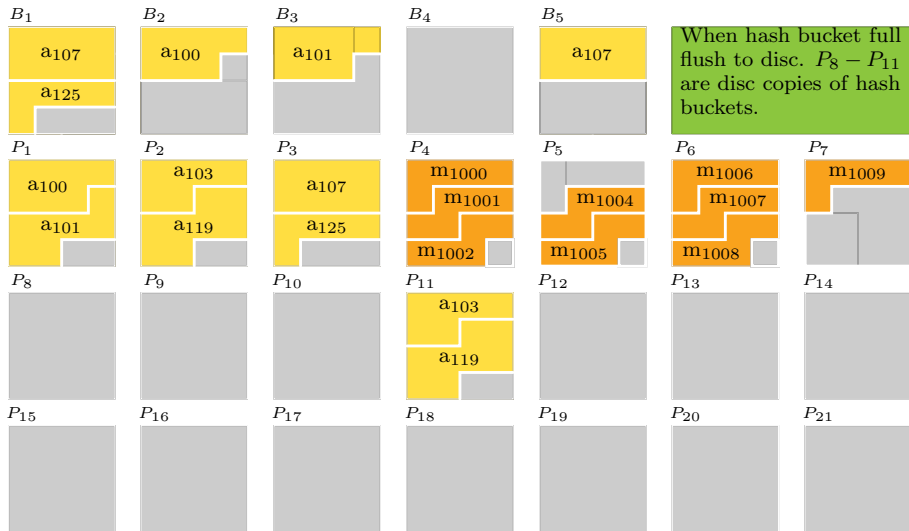
Hash Join



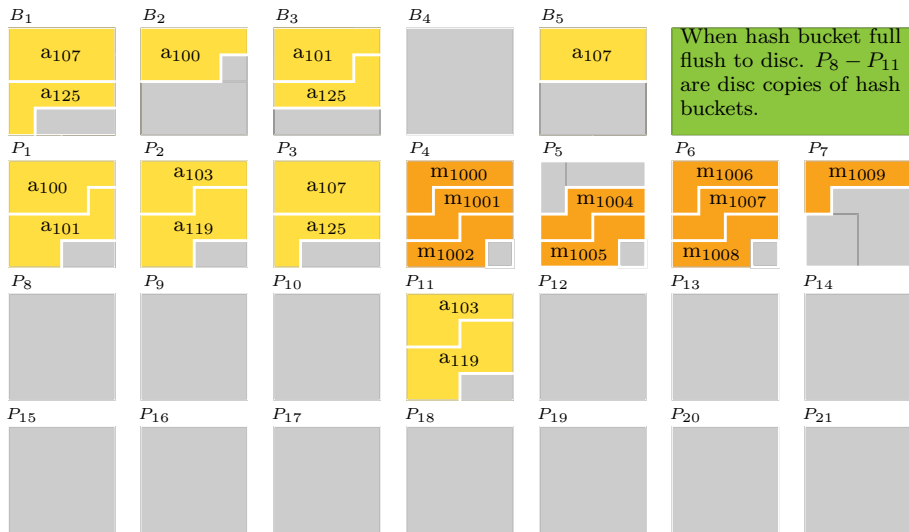
Hash Join



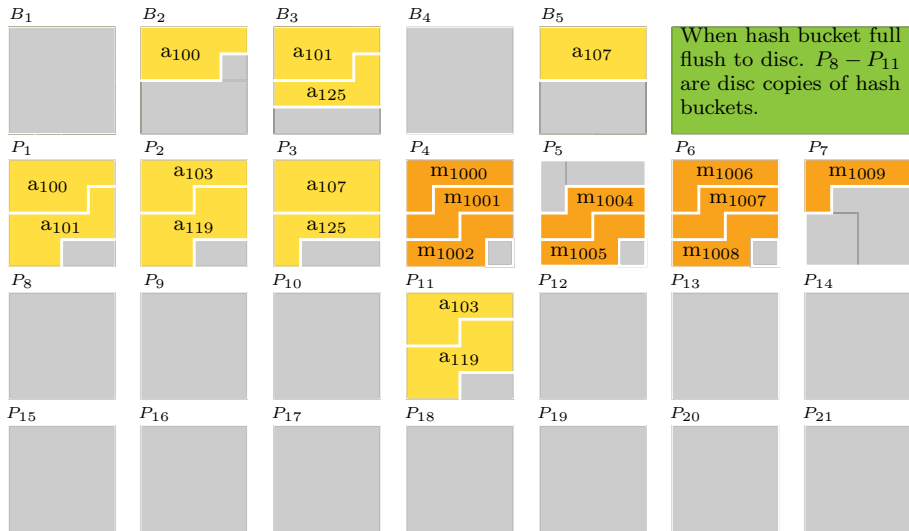
Hash Join



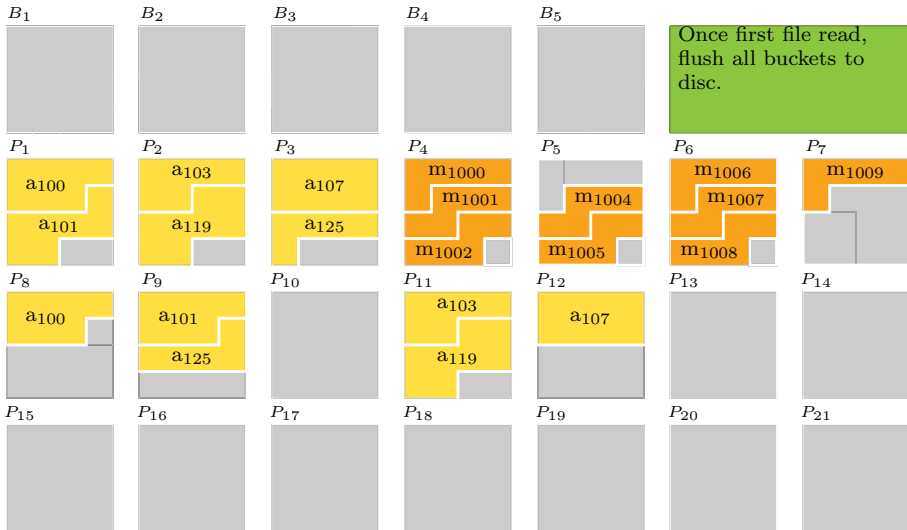
Hash Join



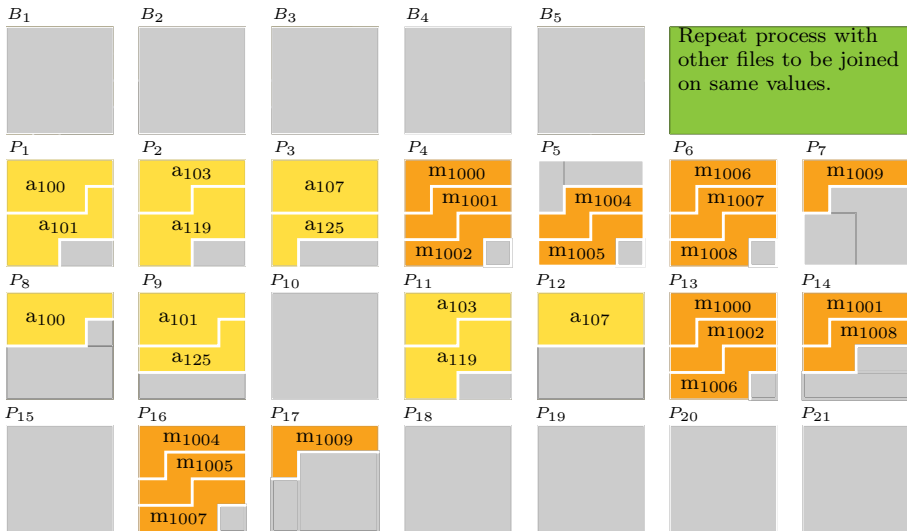
Hash Join



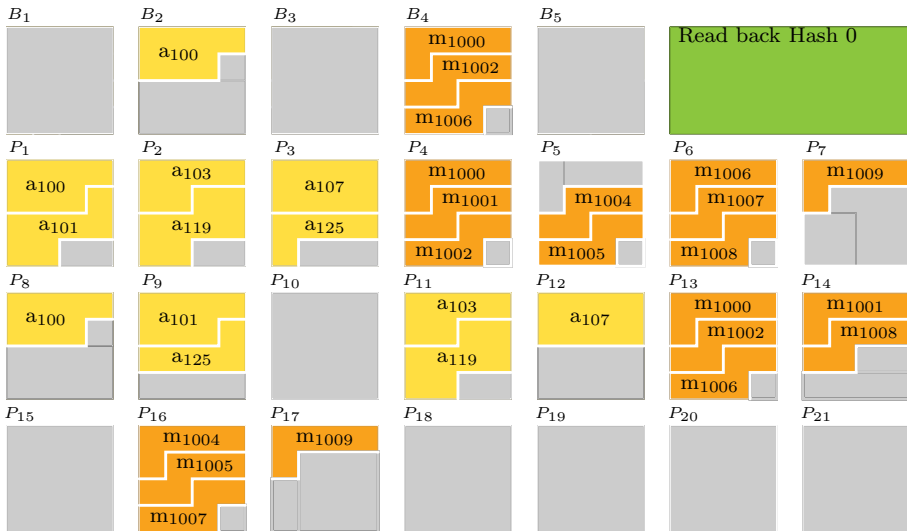
Hash Join



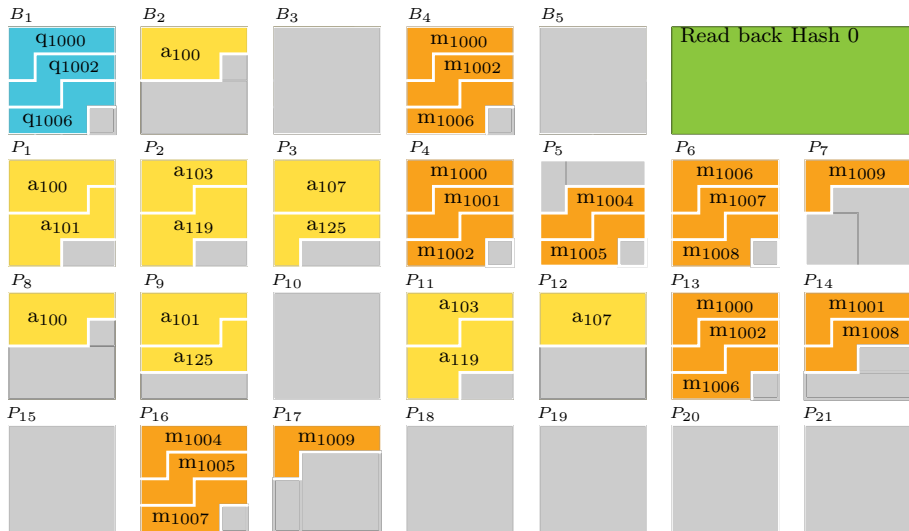
Hash Join



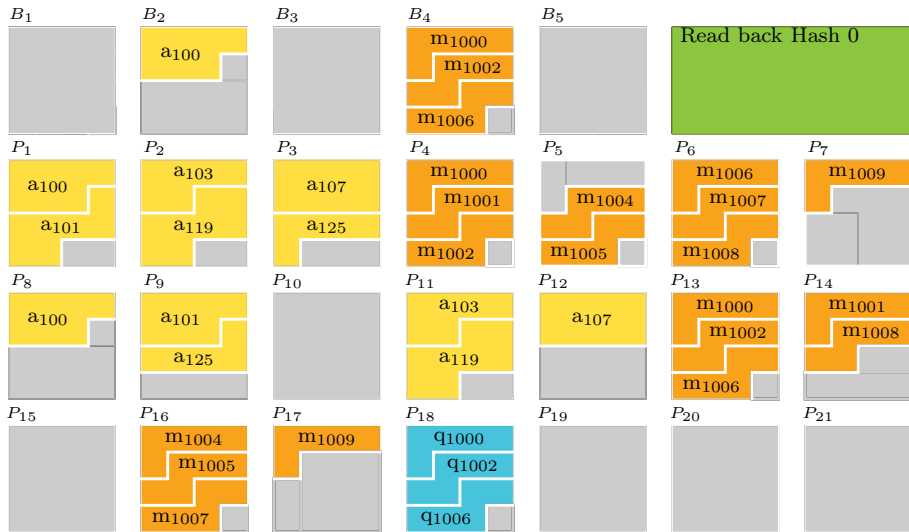
Hash Join



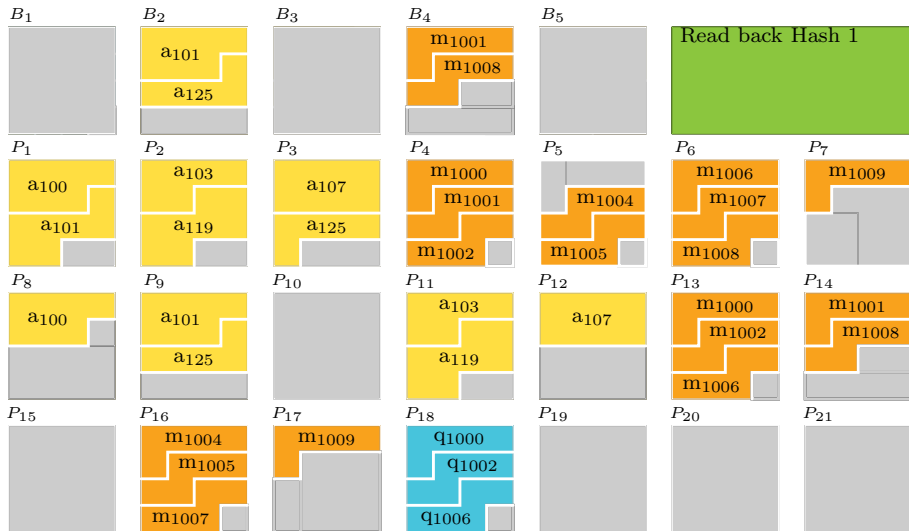
Hash Join



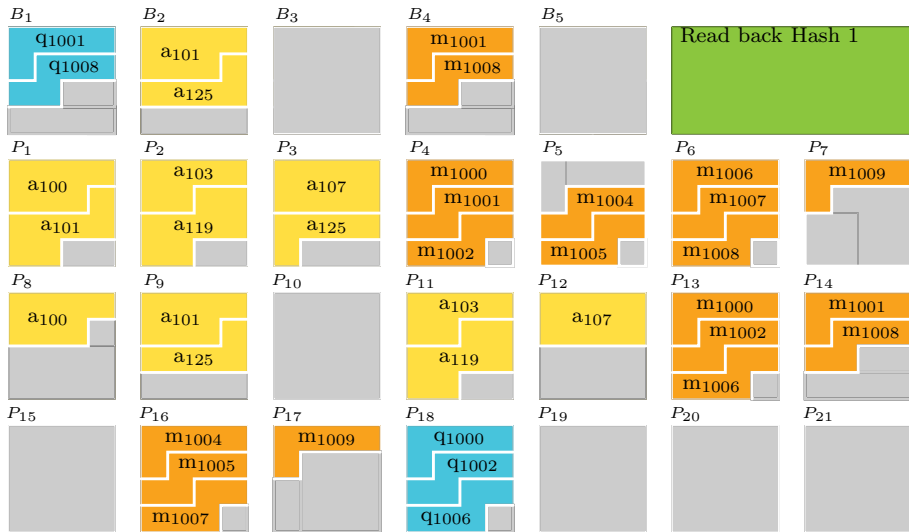
Hash Join



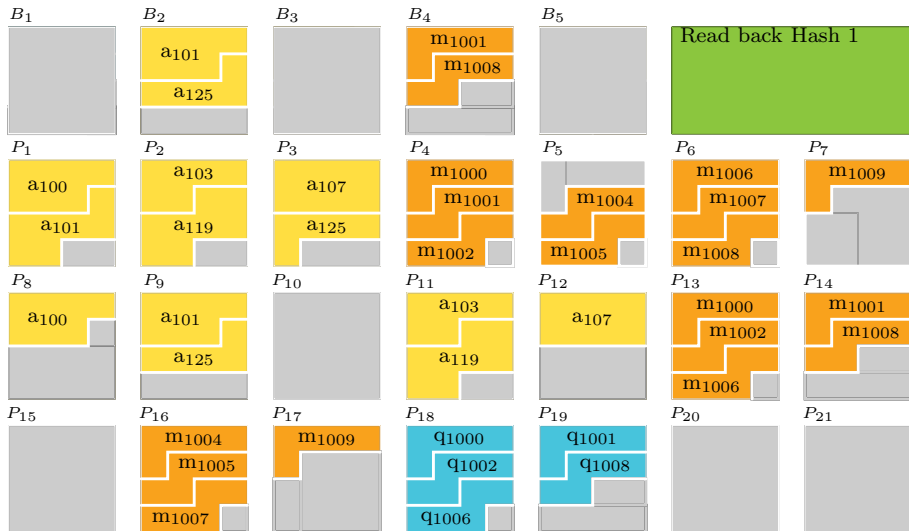
Hash Join



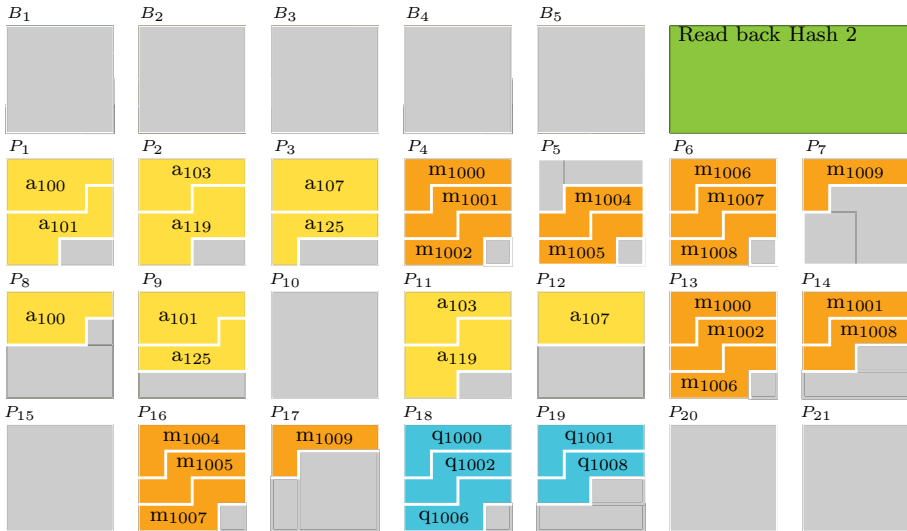
Hash Join



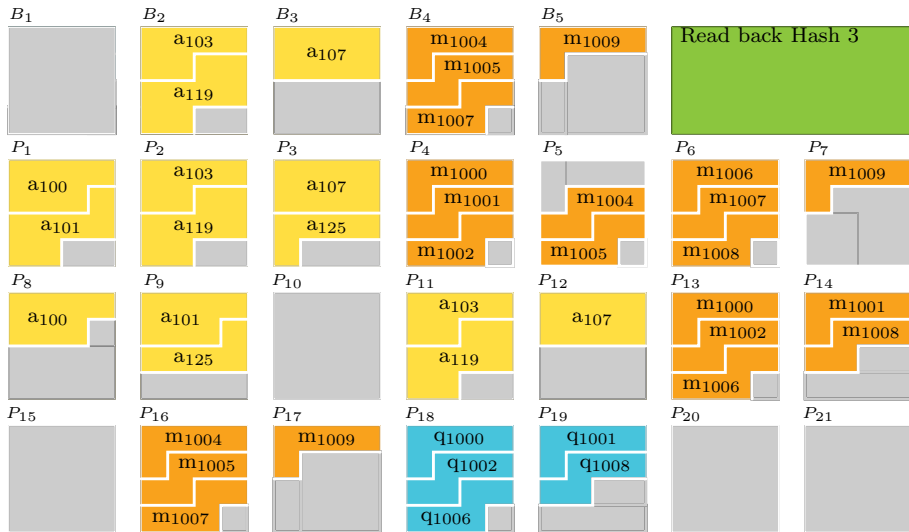
Hash Join



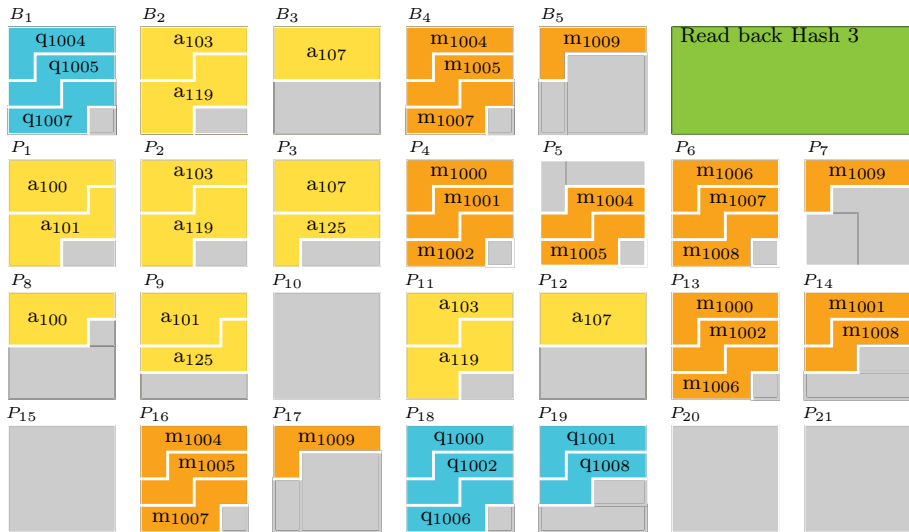
Hash Join



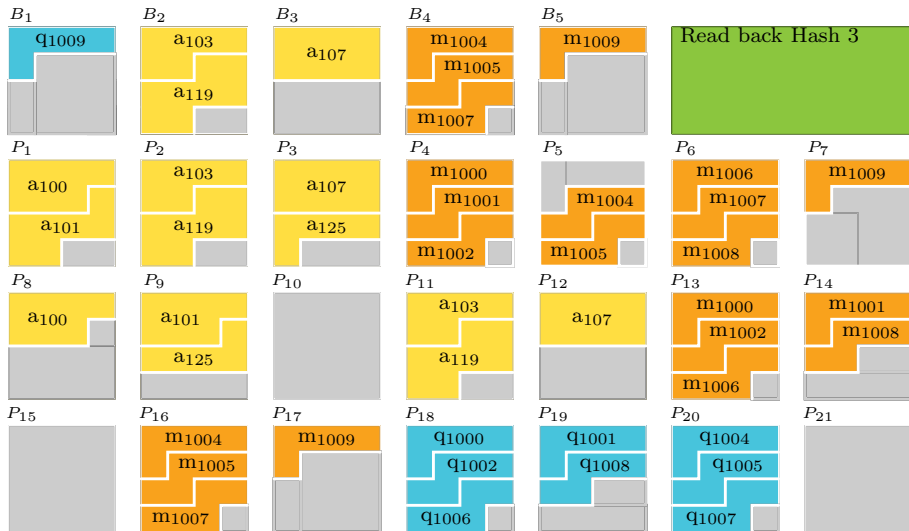
Hash Join



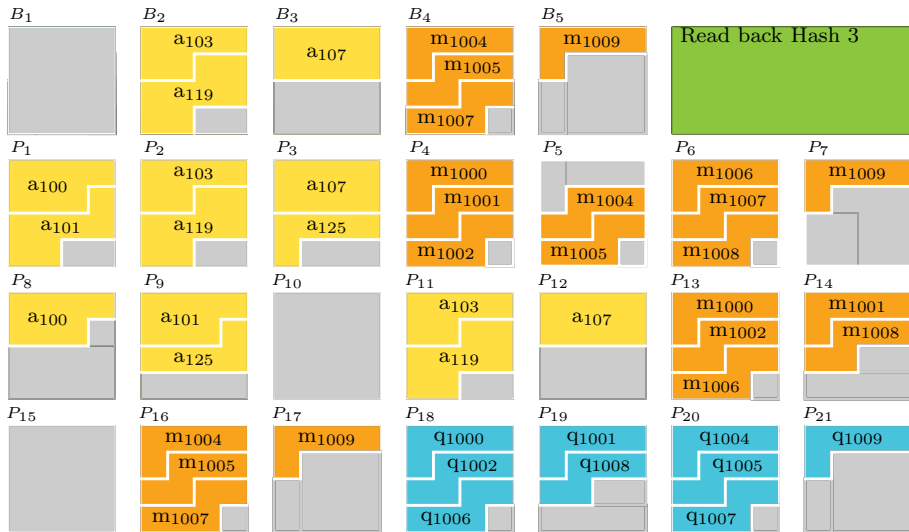
Hash Join



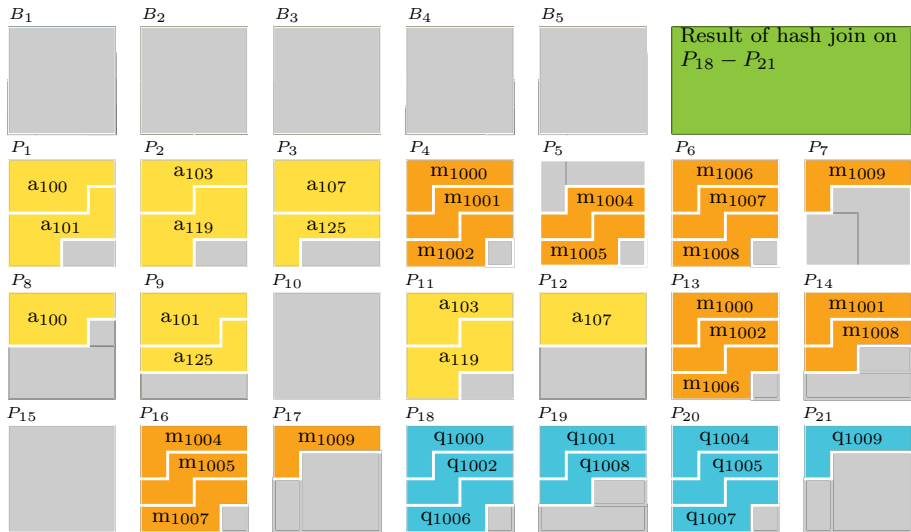
Hash Join



Hash Join



Hash Join



Hash Join: Analysis

$$R \bowtie^{HJ} S$$

- Cost of reading in pages to generate hash buckets

$$|Pages(R)| + |Pages(S)|$$

Hash Join: Analysis

$$R \bowtie^{HJ} S$$

- Cost of reading in pages to generate hash buckets

$$|Pages(R)| + |Pages(S)|$$

- Cost of writing out hash buckets

$$|Pages(R)| + |Pages(S)|$$

Hash Join: Analysis

$$R \bowtie^{HJ} S$$

- Cost of reading in pages to generate hash buckets

$$|Pages(R)| + |Pages(S)|$$

- Cost of writing out hash buckets

$$|Pages(R)| + |Pages(S)|$$

- Cost of reading in hash buckets

$$|Pages(R)| + |Pages(S)|$$

Hash Join: Analysis

$$R \stackrel{HJ}{\bowtie} S$$

- Cost of reading in pages to generate hash buckets

$$|Pages(R)| + |Pages(S)|$$

- Cost of writing out hash buckets

$$|Pages(R)| + |Pages(S)|$$

- Cost of reading in hash buckets

$$|Pages(R)| + |Pages(S)|$$

- Total cost

$$3(|Pages(R)| + |Pages(S)|)$$