

①	Materialization	Pipelining
i)	It is a traditional approach to evaluate multiple operations.	i) It is a modern approach to evaluate multiple operations.
ii)	It uses temporary relations for storing the results of the evaluated operations. So, it needs more temporary files and I/O.	ii) It does not use any temporary relations for storing the results of the evaluated operations.
iii)	It is less efficient as it takes time to generate the query results.	iii) It is more efficient way of query evaluation as it quickly generates the results.
iv)	It does not have any higher requirements for memory buffers for query evaluation.	iv) It requires memory buffers at a high rate for generating outputs. Insufficient memory buffer will cause thrashing.
v)	No thrashing occurs in materialization. Thus in cases, materialization is having better performance.	v) Poor performance if thrashing occurs.
vi)	The overall cost includes the cost of operations plus the cost of reading and writing results on temporary storage.	vi) It optimizes the cost of query evaluation. As it does not include the cost of reading and writing the temporary storage.

② Advantages & i) Reduces the search space for optimization strategy.

- ii) Allows the query optimizer to be based on dynamic processing technique.
- iii) Convenient for Pipelining.

Disadvantages: i) Reduction of search space might skip some low cost evaluation techniques.

- ii) Only fully pipelined strategies are used.
- iii) Many alternative execution strategies are not considered.

3) a) S1: $\sigma_{\text{roomNo}=1 \wedge \text{hotelNo}='4001'}(\text{Room})$

SC = 1 as room no & hotel no as key attribute of room using hash = Cost - 1

$$\text{Linear search cost} = \lceil 10000/2000 \rceil / 2 = 25$$

S2: $\sigma_{\text{type}='D'}(\text{Room})$

$$\begin{aligned} \text{SC}_{\text{type}}(\text{Room}) &= n_{\text{tuples}}(\text{Room}) / n_{\text{Distinct type}}(\text{Room}) \\ &= 10000/10 = 1000 \end{aligned}$$

$$\text{Cost} = 2 + 1000/200 = 5$$

$$\text{Linear search cost} = 50$$

S3: $\sigma_{\text{hotelNo}='4002'}(\text{Room})$

$$\text{SC}_{\text{hotelNo}}(\text{Room}) = 10000/50 = 200$$

$$\text{Cost} = 2 + 200/200 = 3$$

S4: $\sigma_{\text{price} > 100}(\text{Room})$

$$\text{Cost} = 2 + (50/2) + 1000/2 = 5027$$

$$\text{Linear search} = 50$$

S5: $\sigma_{\text{type}='S' \wedge \text{hotelNo}='4003'}(\text{Room})$

Secondary index searches costs are 5 & 3
Linear cost = 50

$J_5: \text{Type} = 'S' \vee \text{Price} \leq 100 (\text{Room})$

Use secondary index on type & price take union
Linear search cost = 50

b) $J_1: (\text{Hotel}) \times \text{hotelNo} (\text{Room})$

Assume NBuffer = 100

Block nested loop = $n\text{Blocks}(\text{Room}) = 10000/200 = 50$

$n\text{Blocks}(\text{Hotel}) = 50/40 = 1$

$R = \text{Hotel} = S = \text{Room}$

If buffer has 1 block for R & S

$$\text{Cost} = n\text{Block}(R) + n\text{Block}(R) \& n\text{Blocks}(S) \\ = 1 + (50 \times 2) = 102$$

If $R(\text{Hotel})$ fits no buffer

$$\text{Cost} = n\text{Block}(R) + n\text{Blocks}(S) = 50 + 2 = 52$$

As Hotel is primary key

$$n\text{ blocks}(R) + n\text{ tuples}(R) \times n\text{ levels hotelNo} + \left(\frac{S \& R}{n\text{ blocks}(R)} \right) \\ 2 + 50 \times (2 + 1) = 2 + 150 = 152$$

Best merge

$$n\text{ blocks } R \times [\log_2(n\text{ blocks}(R))] + n\text{ blocks } S [\log_2 n\text{ blocks}(S)] \\ = 2 \times \log_2(2) + 50 \times [\log_2(50)] \\ = 2 + 250 = 252$$

$$\text{for merge} = 50 + 2$$

Hash

If hash fits into memory

$$3(n\text{ blocks}(R) + n\text{ blocks}(S))$$

$J_6: (\text{Booking}) \times \text{roomNo} (\text{Room})$

$$n\text{Blocks}(\text{Booking}) = 1667$$

$$n\text{Blocks}(\text{Room}) = 50$$

If 1 block is allowed =

$$\begin{aligned} & \cdot n\text{Block}(R) \rightarrow (n\text{Block}(R) * n\text{Block}(S)) \\ & = 50 + 50 * 1667 = 83400 \end{aligned}$$

If room fits in memory:

$$n\text{Blocks}(R) + n\text{Blocks}(S) = 50 + 1667 = 1717$$

Sort Merge

$$\begin{aligned} \text{If unsorted } n\text{Block}(R) [\log_2(n\text{Block}(R))] + n\text{Block}(S) \\ = 50 \log_2 50 + 1667 + \log_2 1667 = 18381 \end{aligned}$$

Hash

$$2(1667 + 50) = 5151$$

C) P1: $\pi_{\text{hotelNo}}(\text{Hotel})$

Using Sorting

$$\begin{aligned} n\text{Blocks}(R) + n\text{Blocks}(R) * \log_2(n\text{Block}(R)) \\ 2 + 2 * \log_2 2 = 4 \end{aligned}$$

Using hashing

As hotel no is pk, cost remains same

P3: $\pi_{\text{price}}(\text{Room})$

Using Sorting

$$\begin{aligned} \text{As price is not pk } [n\text{Block}(R) + n\text{Block}(R) * \log_2(n\text{Block}(R))] \\ = 50 + 50 * \log_2 50 = 350 \end{aligned}$$

Using hashing

$$SC\text{ price}(\text{Room}) = 10000/50 = 200$$

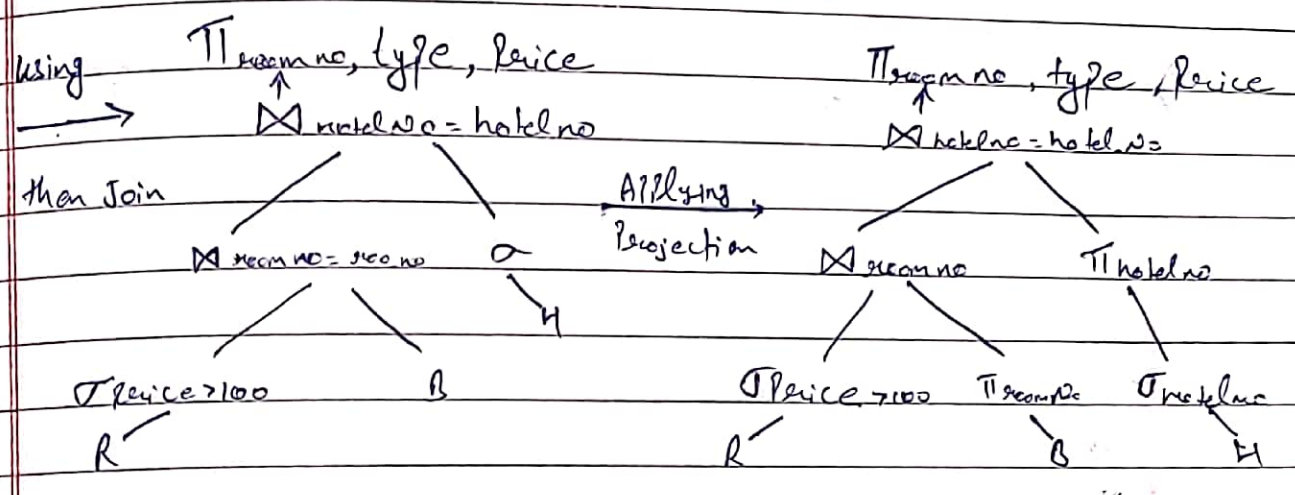
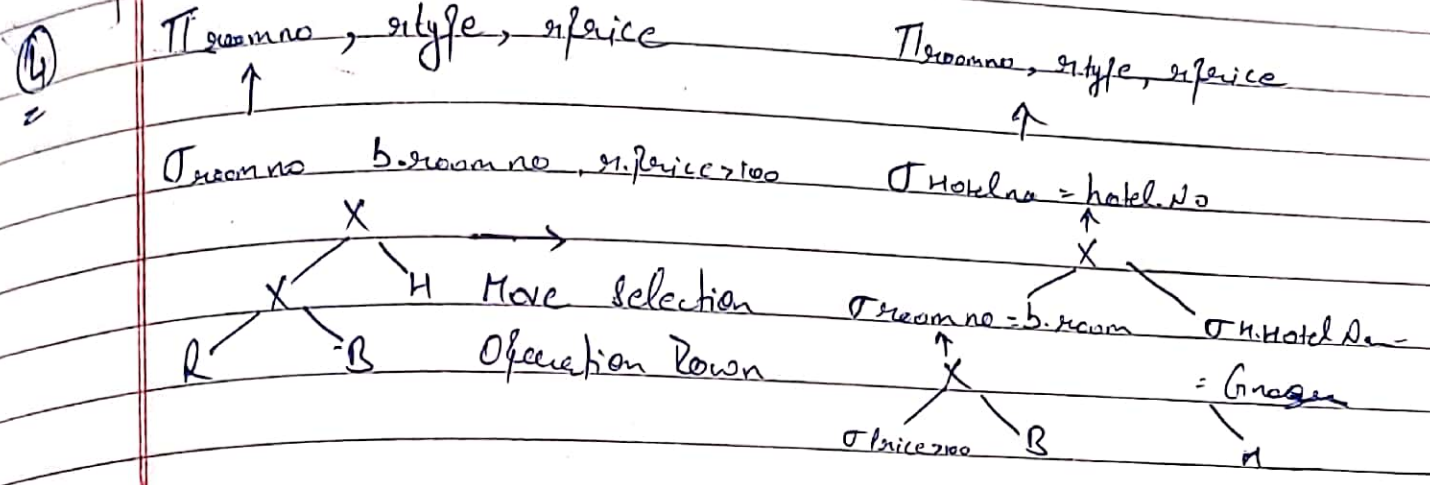
no = 1 (we need 1 block for SC price)

$$n\text{Blocks}(\text{Room}) + nB = 51$$

P5: $\pi_{\text{hotelNo, price}}(\text{Room})$

Using Sorting = 350

using hashing = Hotel type has 51, price no. 51
 So overall 51



⑤ If only left side of join is allowed to be something previous join, it forms left deep tree. Similarly if right side of join is allowed to form previous join, it forms right deep tree. Both these trees are called Linear Algebra trees. Neither of these are Non Linear Algebra trees.

