

Solve the following exercise (you can do it with one of your classmates), and hand in the solution at the beginning of the lecture on "Distributed Query Processing":

Consider the following distributed schema for the Project relation (primary key underlined).

Project(pno, name, head, budget, city):

$$P1 = \sigma_{budget < 100000}(Project)$$

$$P2 = \sigma_{100000 \leq budget \leq 500000}(Project)$$

$$P3 = \sigma_{budget > 500000}(Project)$$



You can assume that this fragmentation strategy is correct (i.e., complete, disjoint and re-constructible). Given the following query:

SELECT * FROM Project WHERE budget > 90000 AND budget < 200000;

Reproduce the data location (i.e., express the query in terms of fragments) and determine which steps the reduction phase would follow.

$$Q = \sigma_{90000 < budget < 200000}$$

$$Q_1 = \sigma_{90000 < budget < 200000} (P_1 \cup P_2 \cup P_3)$$

$$= \sigma_{90000 < budget < 200000} (P_1) \cup \sigma_{90000 < budget < 200000} (P_2) \cup \sigma_{90000 < budget < 200000} (P_3)$$

$$= \sigma_{cond} (\sigma_{budget < 100000}) \cup \sigma_{cond} (\sigma_{100000 \leq budget \leq 500000}) \cup \sigma_{cond} (\sigma_{budget > 500000})$$

$$= \sigma_{budget < 100000 \wedge 90000 < budget \wedge budget < 200000} \rightarrow \text{no null}$$

per simplificar, cond = $90000 < budget < 200000$

$$\sigma_{100000 \leq budget \wedge budget \leq 500000 \wedge 90000 < budget \wedge budget < 200000} \rightarrow \text{no null}$$

$$\sigma_{500000 < budget \wedge 90000 < budget \wedge budget < 200000} \rightarrow \text{nil} \rightarrow \text{no fa falta}$$

$$= \sigma_{90000 \leq budget < 100000} (P) \cup \sigma_{100000 \leq budget < 200000} (P)$$

$$= \sigma_{90000 < budget < 100000} (P_1) \cup \sigma_{budget < 200000} (P_2)$$

Per tant, la nova query serà: $Q_1' = \sigma_{90000 < budget} (P_1) \cup \sigma_{budget < 200000} (P_2)$

