

DISTRIBUTED QUERY PROCESSING

Name: **Marc Duran López.**

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

`Project(pno, name, head, budget, city);`

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

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

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

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

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

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

La consulta original es:

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

Primero, expresamos esta consulta en términos de fragmentos:

$Q = \sigma_{90000 \leq \text{budget} \leq 200000}(\text{Project})$

Esto significa que estamos buscando proyectos cuyo presupuesto esté en el rango de 90,000 a 200,000.

A continuación, expresamos Q en términos de los fragmentos P1, P2 y P3:

$Q' = \sigma_{90000 \leq \text{budget} \leq 200000} (P1 \cup P2 \cup P3)$

Esto se debe a que la consulta se aplica a la relación de proyecto en su totalidad, que está fragmentada en P1, P2 y P3.

Luego, aplicamos la reducción de unión-selección. Esto significa que aplicamos la condición de selección a cada fragmento y luego unimos los resultados:

$$Q' = \sigma_{90000 \leq \text{budget} \leq 200000} (\sigma_{\text{budget} < 100000} (P1)) \cup \sigma_{90000 \leq \text{budget} \leq 200000} (\sigma_{100000 \leq \text{budget} \leq 500000} (P2)) \cup \sigma_{90000 \leq \text{budget} \leq 200000} (\sigma_{\text{budget} > 500000} (P3))$$

Aquí, aplicamos la condición de selección a cada fragmento por separado.

Finalmente, simplificamos aún más la consulta y llegamos a la respuesta final:

$$Q' = \sigma_{90000 \leq \text{budget} < 100000} (P1) \cup \sigma_{100000 \leq \text{budget} < 200000} (P2)$$

Esto se debe a que en el fragmento P1, los proyectos tienen un presupuesto menor a 100,000, por lo que solo necesitamos seleccionar aquellos cuyo presupuesto está en el rango de 90,000 a 100,000. En el fragmento P2, los proyectos tienen un presupuesto en el rango de 100,000 a 500,000, por lo que seleccionamos aquellos cuyo presupuesto está en el rango de 100,000 a 200,000.

Por lo tanto, la respuesta final es:

$$Q' = \sigma_{90000 \leq \text{budget} < 100000} (P1) \cup \sigma_{100000 \leq \text{budget} < 200000} (P2)$$

Esto representa los proyectos que cumplen con la condición de presupuesto en la consulta original después de aplicarla a los fragmentos P1 y P2.