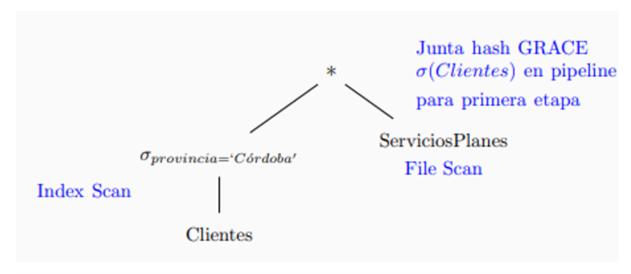
Parcialito 6 - Base de Datos Mariana Juarez Goldemberg - 108441





CLIENTES	SERVICIOSPLANES
n(Clientes) = 2.000.000	n(ServiciosPlanes) = 60.000
B(Clientes) = 200.000	B(ServiciosPlanes) = 6.000
V(provincia, Clientes) = 20	$V(\text{cod_plan, ServiciosPlanes}) = 200$
$V(\text{cod_plan}, \text{Clientes}) = 200$	
H(I(provincia, Clientes)) = 2	

a) Cálculo del costo del plan de ejecución (según cantidad de accesos a disco)

01

La primera operación será la selección desde la tabla Clientes de los clientes cuya provincia sea Córdoba. Se utilizará un índice de *clustering* por el atributo provincia para acceder a la tabla Clientes.

Costo(O1) = H(I(provincia, Clientes)) + B(Clientes) / V(provincia, Clientes)

Siendo H(I(provincia, Clientes)) = 2, B(Clientes) = 200000, V(provincia, Clientes) = 20

Costo(01) = 2 + 200000 / 20 = 10002

Estimo la cardinalidad de la 01 como:

Mariana Juarez Goldemberg - 108441



n(01) = n(Clientes) / V(provincia, Clientes) = 2000000 / 20 = 100000

Y la cantidad de bloques devueltos por el operador será, calculando F(Clientes) = n(Clientes) / B(Clientes) = 10:

$$B(01) = n(01) / F(Clientes) = 100000 / 10 = 10000$$

02

La segunda operación a realizar será la Junta por Hash Grace en pipeline

Costo(02) = 2 * B(01) + 3 * B(ServiciosPlanes) = 2 * 10000 + 3 * 6000 = 38000

Costo(Total) = Costo(01) + Costo(02) = 10002 + 38000 = 48002

b) Estimación de la cardinalidad del resultado de la junta

Para estimar la cardinalidad asumo que tengo una distribución equitativa en 01.

 $V(\text{cod_plan}, 01) = 200 / 20 = 10$ (se reparten equitativamente los códigos de plan entre las provincias)

 $n(02) = n(01) * n(ServiciosPlanes) / máx(V(cod_plan, 01), V(cod_plan, ServiciosPlanes))$ n(02) = 100000 * 60000 / máx(10,200) = 30000000

1/F(02) = 1/F(Clientes) + 1/F(ServiciosPlanes) = 1/10 + 1/10 = 1/5F(02) = 5

F(ServiciosPlanes) = n(ServicosPlanes) / B(ServiciosPlanes) = 10

B(02) = n(02) / F(02) = 30000000 / 5 = 6000000