C10

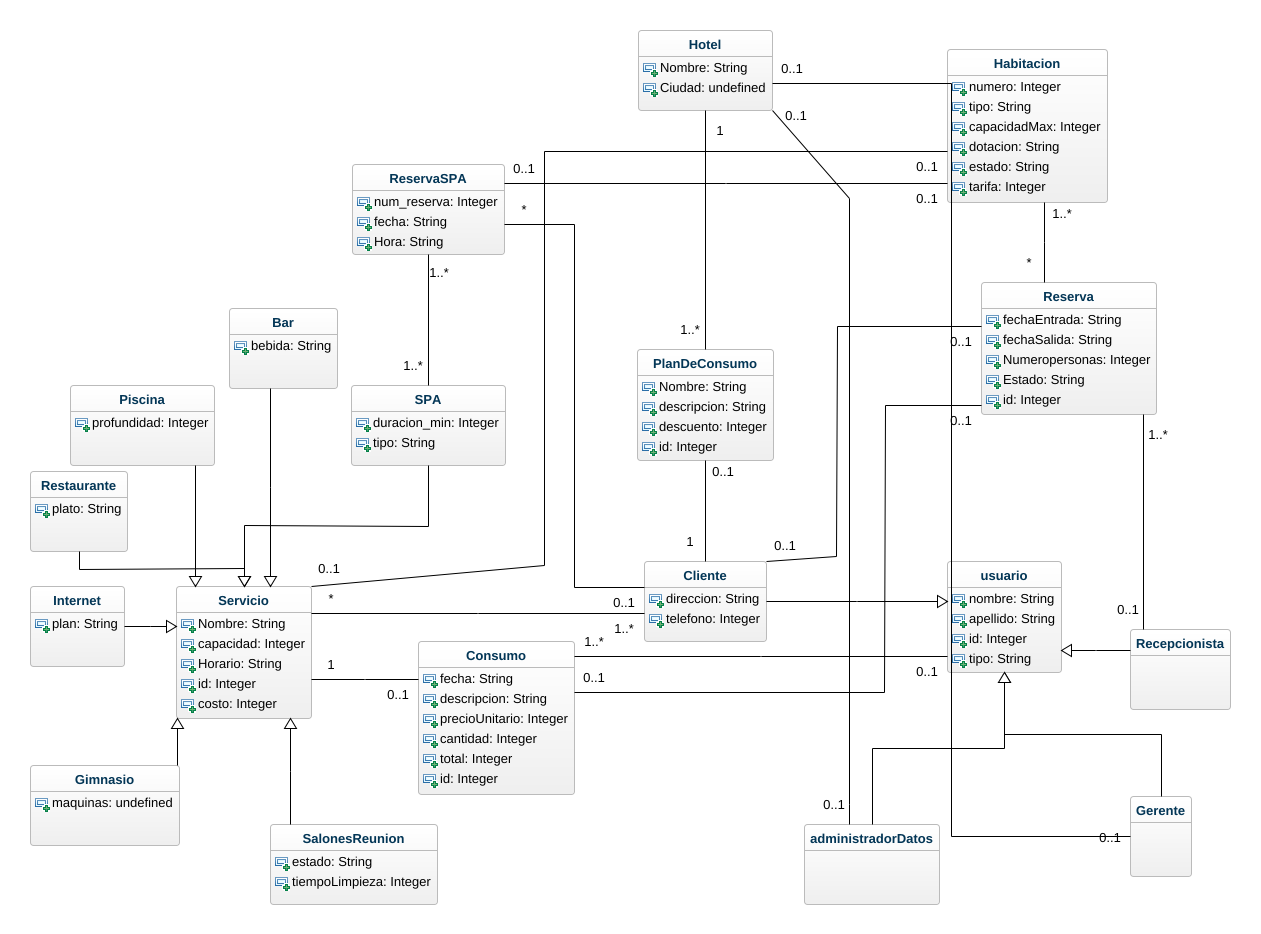
Andres Serrano

Samuel Ramirez

Juana Mejia

Analisis

Al analizar el impacto que representa la introducción de los nuevos atributos se encontró que en el diseño inicial falto la clase de salones de reunión/conferencia, estos se volvieron una clase ya que lo único que los diferenciaba son la capacidad que tiene y el tiempo de limpieza luego de uso. El resto de clases y relaciones se mantuvieron igual debido a que no eran afectados por los nuevos requerimientos.



Diseño de Aplicación

Indices

1. Para el requerimiento funcional de “mostrar el dinero recolectado por servicios en cada habitación en el último año corrido” se podría crear un índice sobre las fechas de servicio para así poder encontrar los resultados que se cumplen con el rango presentado de un año y mejorar significativamente el rendimiento. También se podría poner un índice sobre los id de las habitaciones ya que se optimizaría el agrupamiento este atributo. Al ser un índice sobre el id de la habitación, esto implica que habrá alta selectividad dependiendo de la cantidad de cuartos que haya en el hotel. El tipo de índice para los mencionados anteriormente puede ser B+
2. EN el siguiente requerimiento “MOSTRAR LOS 20 SERVICIOS MÁS POPULARES.” Se puede crear un índice sobre la columna de los servicios, esto se debe a que con este se puede optimizar el rendimiento de la consulta y se acelerara la clasificación y recuento de los servicios. En términos de selectividad, como se tienen tantos servicios diferentes, este tiene un alto porcentaje y puede permitir que menos registros sean involucrados en la operación. Este índice puede ser de tipo B+
3. MOSTRAR EL ÍNDICE DE OCUPACIÓN DE CADA UNA DE LAS HABITACIONES DEL HOTEL: para este requerimiento se puede crear sobre los id de las habitaciones, esto ayudara a mejorar el rendimiento para la operación de agrupamiento por id de habitación y calcular el índice de ocupación de manera más rápida. El tipo de índice seria B+.
4. MOSTRAR LOS SERVICIOS QUE CUMPLEN CON CIERTA CARACTERÍSTICA: en este caso se deben identificar las consultas realizadas con mayor frecuencia, se podría hacer un índice para los precios de los servicios para así buscar en un rango, también se podría hacer un índice para la fecha de consumo que permitirá definir si hace parte del rango. Estos dos índices tendrán un alto nivel de selectividad debido a que hay una gran cantidad de precios y fechas por lo que la selectividad será bastante alta y permitirá reducir la cantidad de registros en la consulta. Estos índices podrán ser de tipo B+. Otro índice podría ser sobre empleado id, para definir cual empleado hizo el registro del servicio. Este puede no beneficiar el tiempo de la transacción debido a que pueden haber pocos empleados por lo que la selectividad será baja.
5. MOSTRAR EL CONSUMO EN HOTELANDES POR UN USUARIO DADO, EN UN RANGO DE FECHAS INDICADO: para este requerimiento se debe encontrar el consumo de un usuario en un rango de fechas, por esto sería beneficioso hacer un índice en la fecha de servicio para poder revisar si hace parte del rango, también se podría hacer un índice sobre el id del usuario que consumió el servicio. Estos dos índices tienen un alto nivel de selectividad
6. ANALIZAR LA OPERACIÓN DE HOTELANDES: Para este requerimiento se pueden crear los índices en la fecha de servicio y en el id de las habitaciones, estos permiten identificar rápidamente los días en los que el hotel tuvo la mayor ocupación. La mayoría del análisis implica encontrar fechas y al tener un índice sobre estas, se podrán hacer las consultas más rápido y optimo. Debido a que las fechas son la mayoría de los resultados para analizar la operación de hotelandes, es de suma importancia que sea un proceso rápido, esto se puede hacer en un árbol b+.
7. ENCONTRAR LOS BUENOS CLIENTES: Debido a que se deben conocer el número de días de la estadía y el costo consumido, es importante tener índices sobre la columna de clientes, para identificarlos más fácil, también un índice para fecha de servicio y para los consumos de clientes. Estos índices tienen altos niveles de selectividad debido a la cantidad de valores que pueden tomar y el hecho de que el hotel tiene bastantes clientes diferentes. Pueden ser de tipo B+
8. ENCONTRAR LOS SERVICIOS QUE NO TIENEN MUCHA DEMANDA: Para encontrar estos servicios se debe tener un índice sobre la columna de fecha del servicio debido a que se debe saber cuándo fue tomado el servicio para así determinar si fue utilizado menos de 3 veces. Al ser una fecha, habrán muchísimas posibilidades y permitirá que la selectividad sea de un porcentaje alto. Este índice puede ser de tipo B+.
9. CONSULTAR CONSUMO EN HOTELANDES: Se necesitarían los índices en el servicio para encontrar rápidamente servicios específicos, uno para el id del cliente y uno para la fecha del servicio. Estos tres índices tienen una alta selectividad debido a que pueden tomar demasiados valores distintos. Pueden ser de árbol B+.
10. CONSULTAR CONSUMO EN HOTELANDES – RFC9-V2: para este requerimiento se debe tener un índice para el id del servicio, un índice para el id del cliente que toma el servicio, uno para la columna de la fecha de servicio. Todos estos para poder encontrar la información de los clientes que no han consumido un servicio determinado
11. CONSULTAR FUNCIONAMIENTO: se debe hacer un índice para la fecha de servicio, para el id de servicio y el numero de la habitación. Estos serán índices con alta selectividad debido a que existen muchos servicios, fechas y numero de habitaciones. Se podrá usar árbol B+
12. CONSULTAR LOS CLIENTES EXCELENTES: Para realizar esta consulta se podrían necesitar índices en el id del cliente, la fecha del checkin, checkout, costo de servicios y de id de servicio. Estos índices van a ayudar a filtrar las condiciones necesarias para encontrar a los clientes “excelentes”. Estos índices serán de tipo B+

Carga de Datos

La carga de datos para esta iteración es un reto bastante importante, ya que el volumen que se solicita es alto. Se solicitan 750000 tuplas en las tablas. Claro que, este volumen no se logra en todas las tuplas. Ya que en ciertas tablas no tiene sentido alguno implementar 750000 tuplas, porque no se ajusta al modelo de la vida real. Estas tablas son: Tipo\_usuario, Tipo\_habitacion, Servicios, Bar, Restaurante, Spa, Empleado y Habitación. En las demás tablas si se logró llegar a tal cifra de 750000 tuplas. A continuación vamos a explicar cómo.

Insertar tupla por tupla es un proceso bastante demorado en el que no solo nos demoraríamos pensando cada valor individual. Aun así, si tuviéramos las 750000 sentencias sql, ejecutar un archivo de tal tamaño sentencia por sentencia es muy demorado. Por lo que para la creación de los datos nos apoyamos en el lenguaje Python con el editor VSCO.

En este creamos un archivo el cual nos iba permitir crear archivos csv. Estos archivos csv son una manera fácil de importar datos de gran cantidad en sqldeveloper. El archivo que se utilizó será adjuntado en el repositorio. Pero explicaremos su funcionamiento de manera breve.

En total se crearon 16 archivos CSV, ya que se tenían 16 tablas y por cada tabla se creaba un archivo CSV. Cada tabla poseía una lista con los campos. Luego para crear cada una de las tuplas se hizo un for. Ese for iba desde 1 hasta 750000. Por cada iteración se generaba un dato aleatorio (esto teniendo en cuenta el tipo de dato) y se anexaba a una lista que contenía una tupla de datos (la tupla varía dependiendo de la tabla) y cada tupla de datos era anexada a una lista de tuplas, ya con toda la lista de tuplas se utilizó la función para crear, abrir y escribir en un csv para crear los 16 CSV. Ya luego de tener los 16 archivos CSV creados, en sql developer. Por cada tabla había la opción de importar los datos. Se importaron los datos y finalmente quedaron en la base de datos.

Diseño de las consultas

**RFC2 - MOSTRAR LOS 20 SERVICIOS MÁS POPULARES.**

SELECT SERVICIO\_ID, COUNT(\*) AS CANT\_CONSUMOS

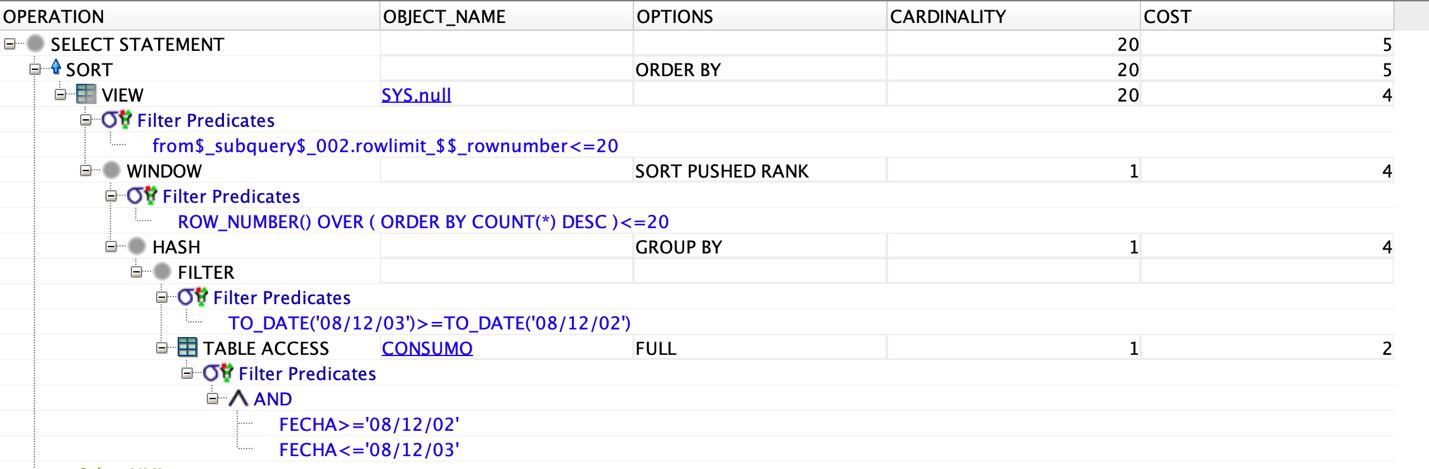
FROM CONSUMO

WHERE FECHA BETWEEN '08/12/02' AND '08/12/03'

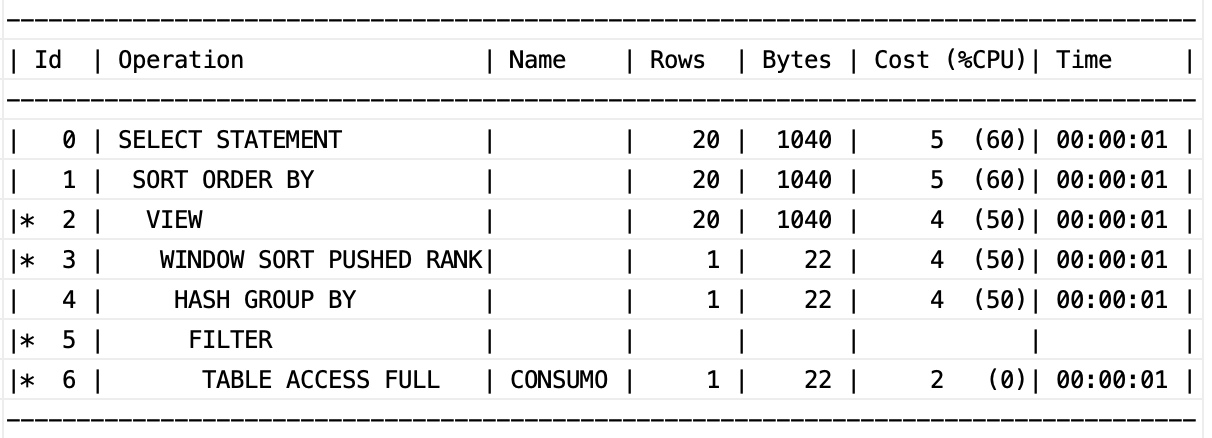
GROUP BY SERVICIO\_ID

ORDER BY CANT\_CONSUMOS DESC;

Plan de ejecucion



Tiempo de la consulta:



**RFC3 - MOSTRAR EL ÍNDICE DE OCUPACIÓN DE CADA UNA DE LAS HABITACIONES DEL HOTEL**

SELECT Reservanhabitaciones.habitacion\_numero,((COUNT(FECHASALIDA-FECHAENTRADA))/365)

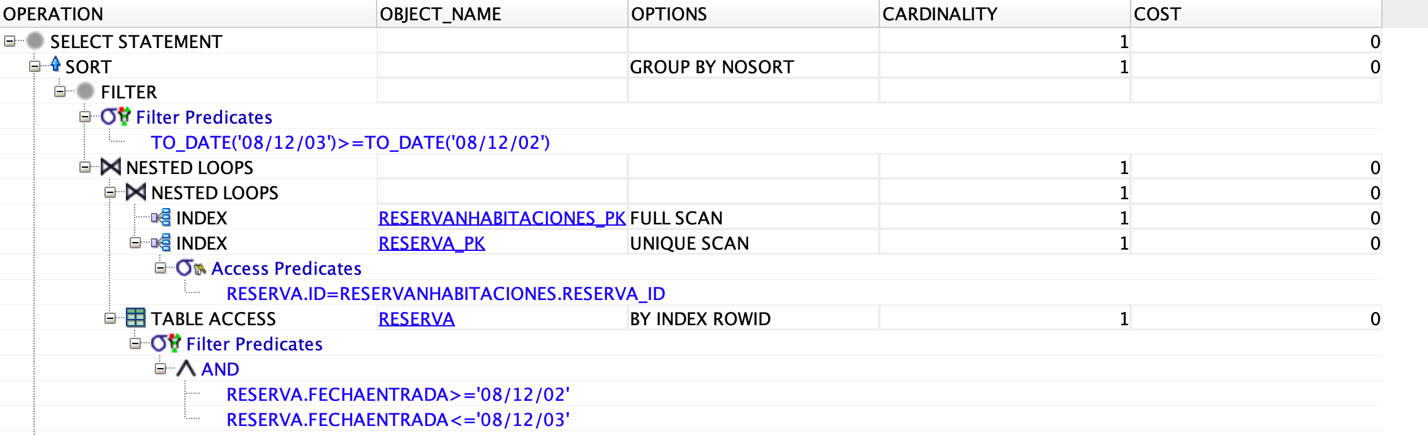
FROM RESERVA

INNER JOIN RESERVANHABITACIONES ON (RESERVA.ID = RESERVANHABITACIONES.RESERVA\_ID)

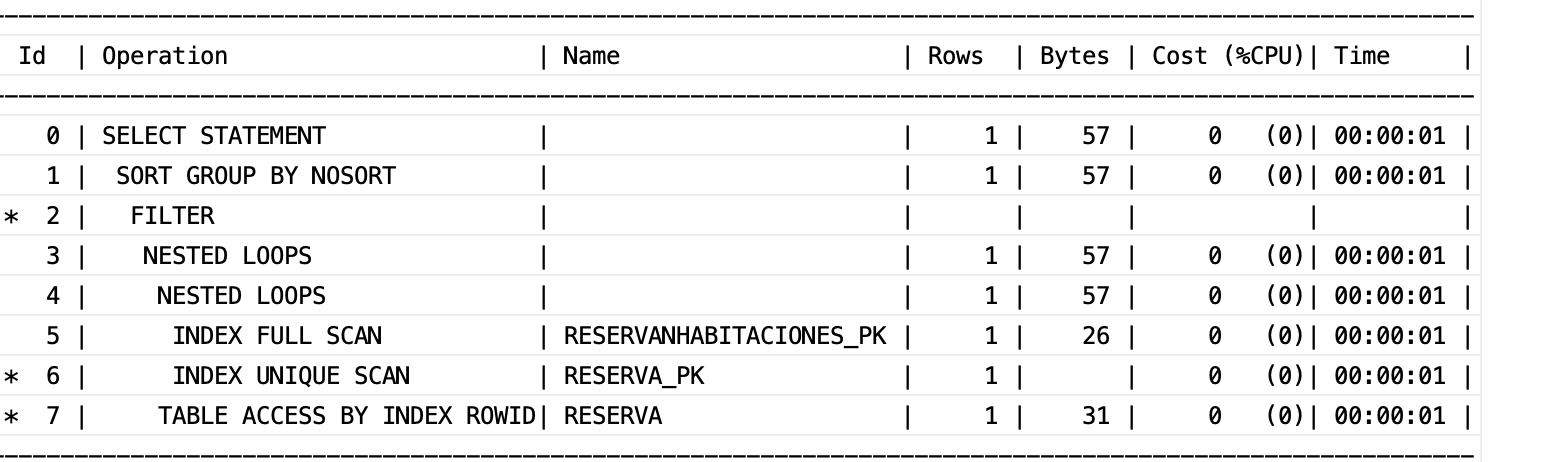
WHERE reserva.fechaentrada BETWEEN '08/12/02' AND '08/12/03'

GROUP BY Reservanhabitaciones.habitacion\_numero;

Plan de ejecucion



Tiempo de ejecución:



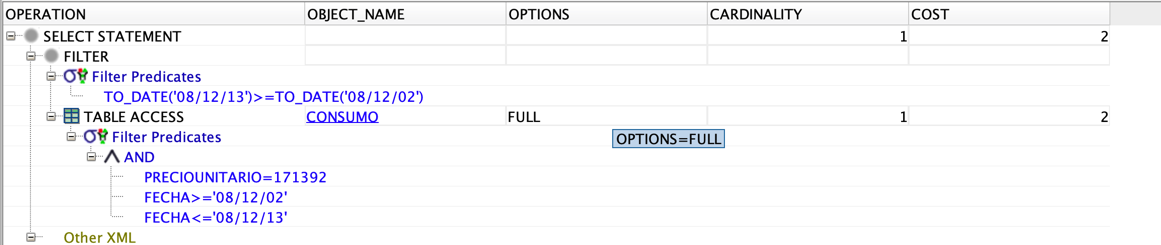
**RFC4 - MOSTRAR LOS SERVICIOS QUE CUMPLEN CON CIERTA CARACTERÍSTICA**

SELECT SERVICIO\_ID

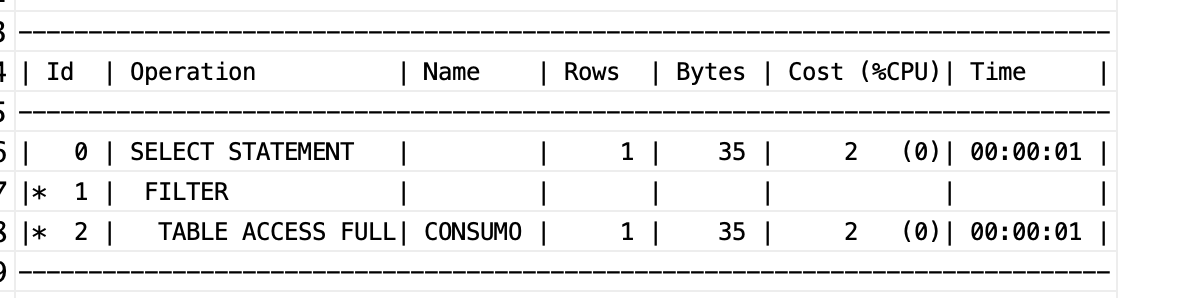
FROM CONSUMO

WHERE (FECHA BETWEEN '08/12/02' AND '08/12/13') AND (PRECIOUNITARIO = '171392');

Plan de ejecucion



Tiempo de ejecusion:



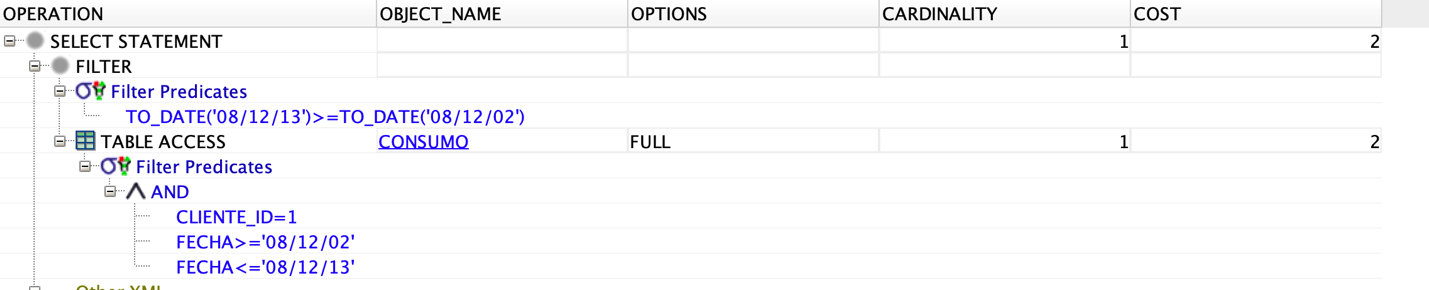
**RFC5 - MOSTRAR EL CONSUMO EN HOTELANDES POR UN USUARIO DADO, EN UN RANGO DE FECHAS INDICADO.**

SELECT ID, DESCRIPCION, TOTAL, CLIENTE\_ID, SERVICIO\_ID

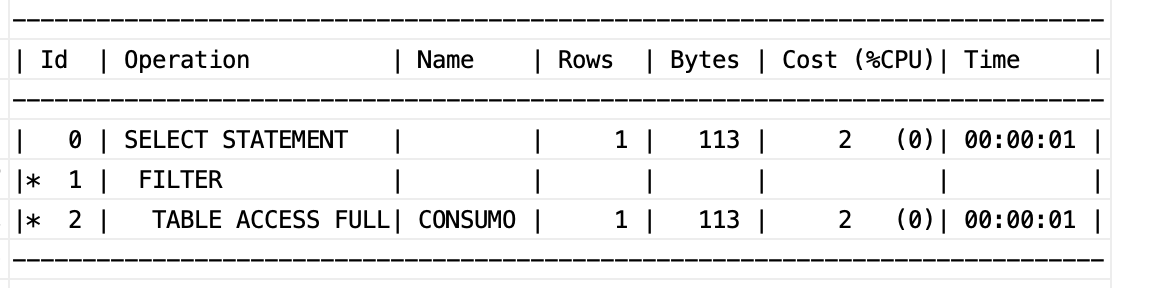
FROM CONSUMO

WHERE (CLIENTE\_ID = 1) AND FECHA BETWEEN '08/12/02' AND '08/12/13';

Plan de ejecusion



Tiempo de ejecución



**RFC6 - ANALIZAR LA OPERACIÓN DE HOTELANDES**

//1

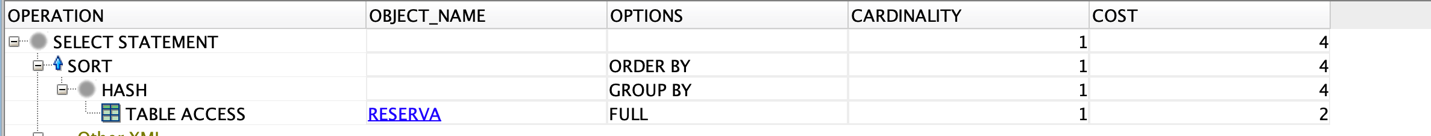
SELECT FECHAENTRADA, SUM(NUMEROPERSONAS) AS TOTAL\_PERSONAS

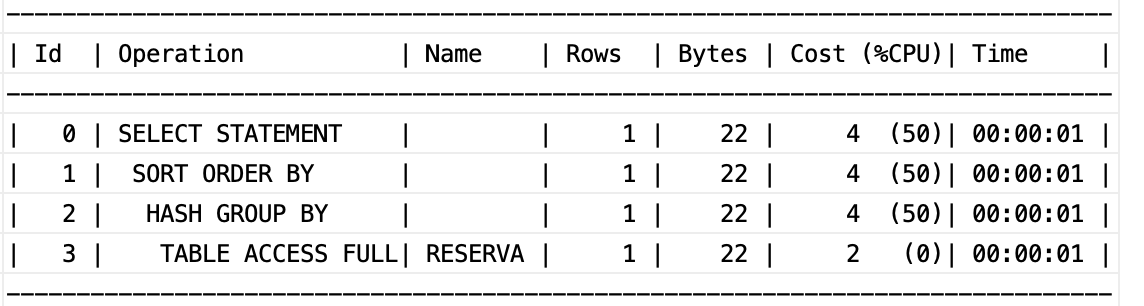
FROM reserva

GROUP BY FECHAENTRADA

ORDER BY TOTAL\_PERSONAS DESC;

Plan de ejecución:



Tiempo:

//2

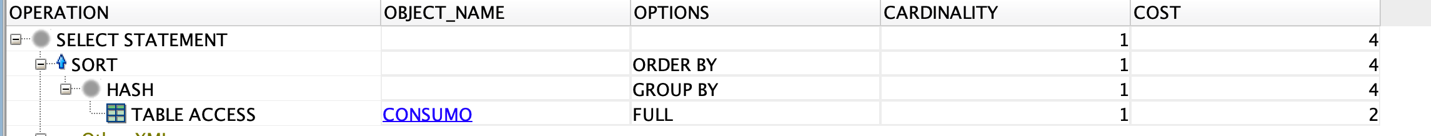
SELECT FECHA, SUM(TOTAL) AS TOTAL\_INGRESOS

FROM consumo

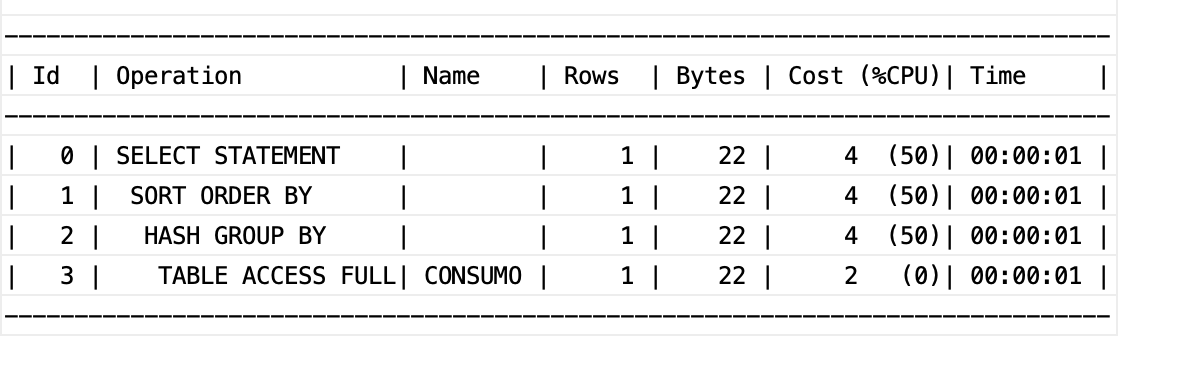
GROUP BY FECHA

ORDER BY TOTAL\_INGRESOS DESC;

Plan de ejecución :



Tiempo:



//3

SELECT FECHAENTRADA, SUM(NUMEROPERSONAS) AS TOTAL\_PERSONAS

FROM reserva

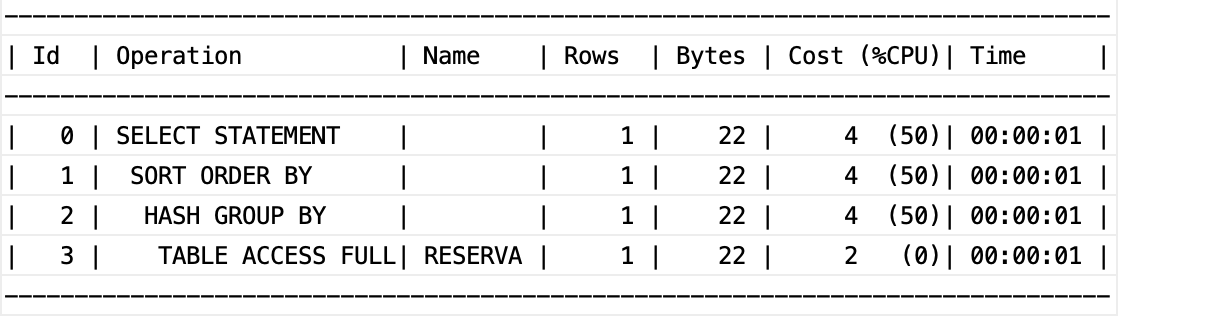
GROUP BY FECHAENTRADA

ORDER BY TOTAL\_PERSONAS Asc;

Plan de ejecución:



Tiempo de ejecución:



**RFC7 - ENCONTRAR LOS BUENOS CLIENTES**

SELECT c.CLIENTE\_ID, SUM(DISTINCT r.FECHASALIDA - r.FECHAENTRADA) AS TOTAL\_DIAS,SUM(c.TOTAL) AS TOTAL\_CONSUMO

FROM reserva r

JOIN consumo c ON r.ID = c.RESERVA\_ID

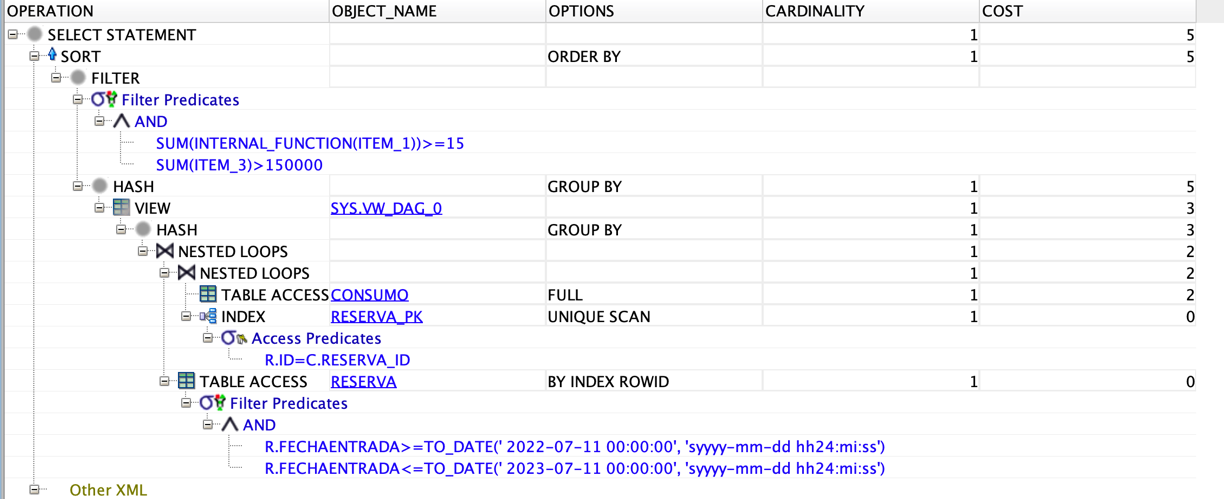
WHERE r.FECHAENTRADA BETWEEN '11/7/2022' AND '11/7/2023'

GROUP BY c.CLIENTE\_ID

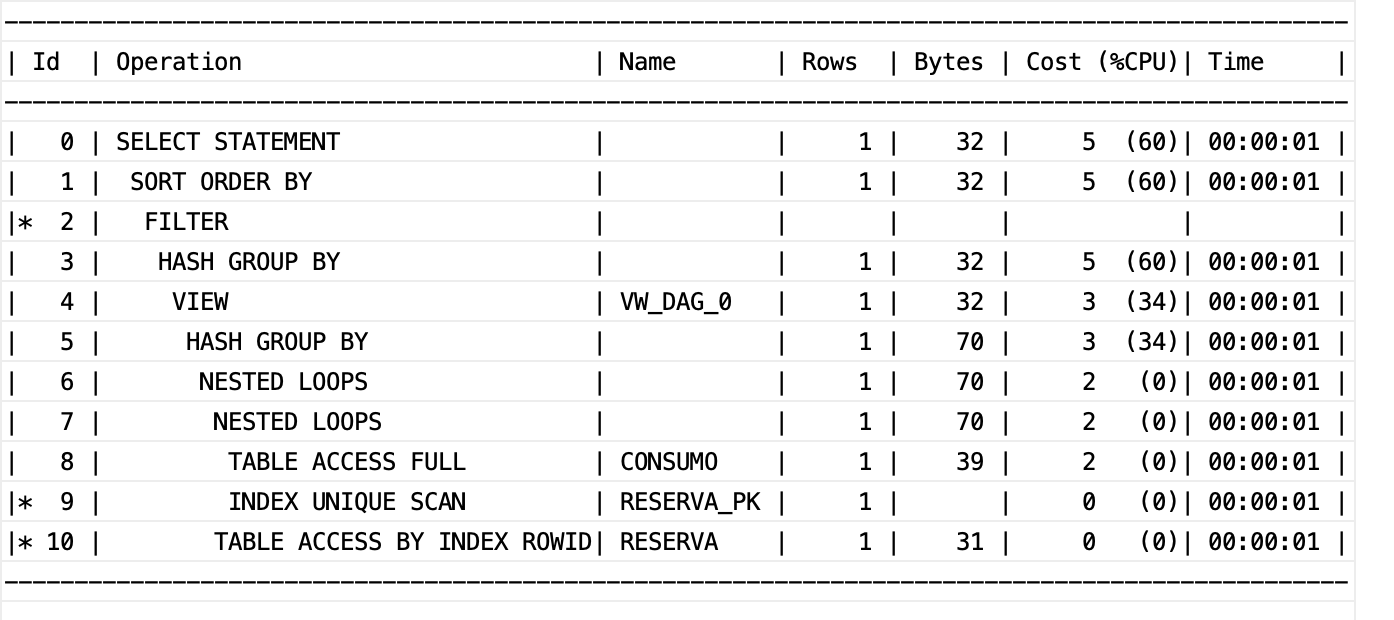
HAVING SUM(DISTINCT r.FECHASALIDA - r.FECHAENTRADA)>=15 AND SUM(c.TOTAL)> 150000

ORDER BY TOTAL\_CONSUMO DESC, TOTAL\_DIAS DESC;

Plan ejecución



Tiempo Ejecución



**RFC8 - ENCONTRAR LOS SERVICIOS QUE NO TIENEN MUCHA DEMANDA**

SELECT SERVICIO\_ID, COUNT(\*) AS NUMERO\_VECES\_SOLICITADO

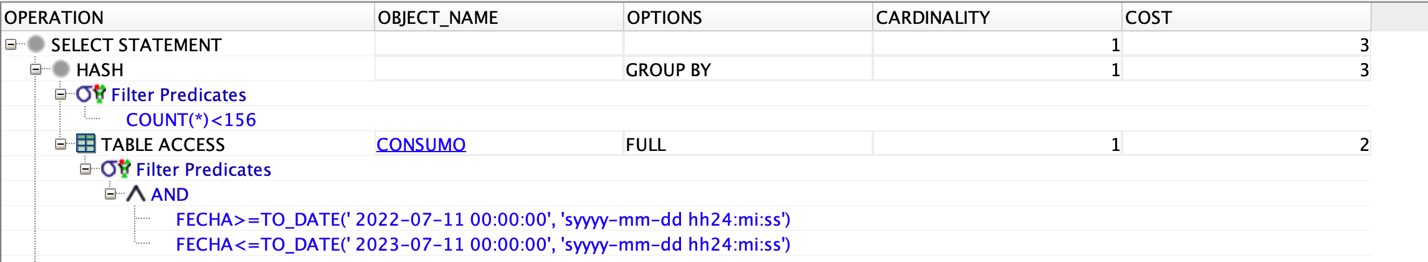
FROM consumo

WHERE FECHA BETWEEN '11/7/2022' AND '11/7/2023'

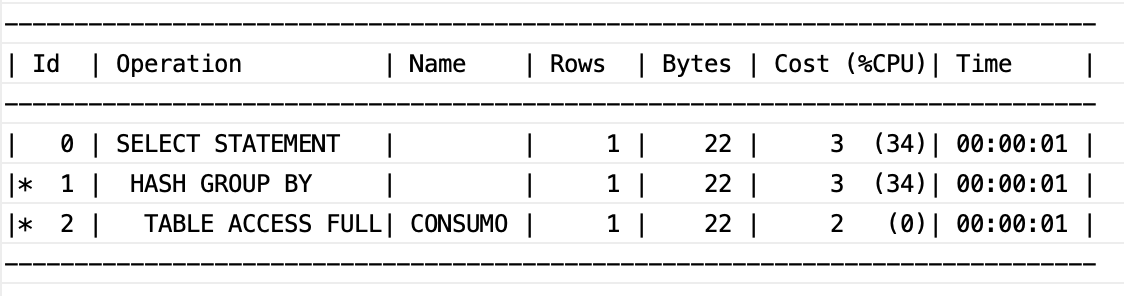
GROUP BY SERVICIO\_ID

HAVING COUNT(\*) < 3 \* 52;

Plan de ejecución:



Tiempo de Ejecución:



**RFC9 - CONSULTAR CONSUMO EN HOTELANDES**

SELECT c.ID AS CLIENTE\_ID,c.DIRECCION, c.TELEFONO, COUNT(\*) AS NUMERO\_CONSUMOS

FROM cliente c

LEFT JOIN consumo co ON c.ID = co.CLIENTE\_ID

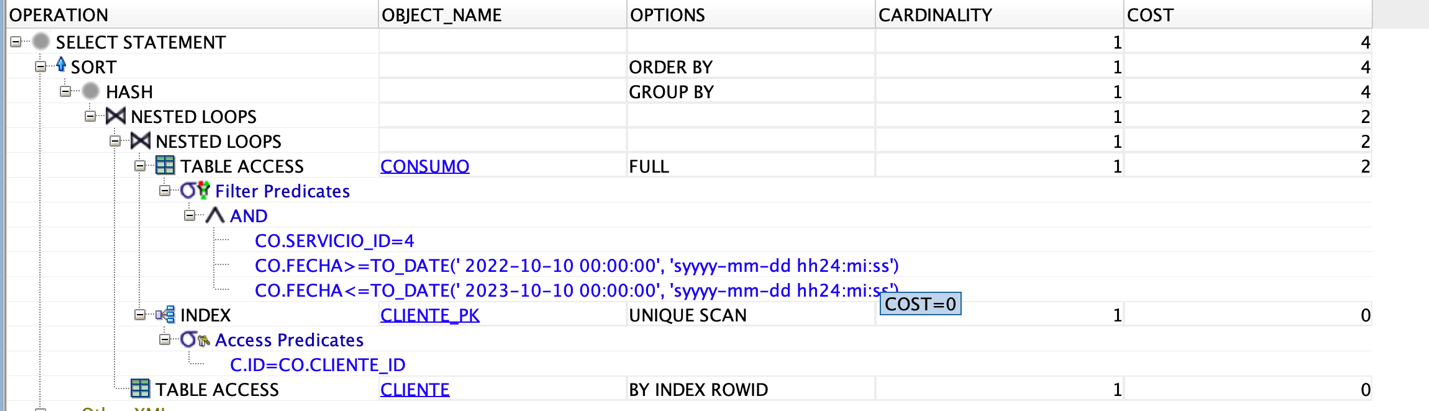
WHERE co.SERVICIO\_ID = 4

AND co.FECHA BETWEEN '10/10/2022' AND '10/10/2023'

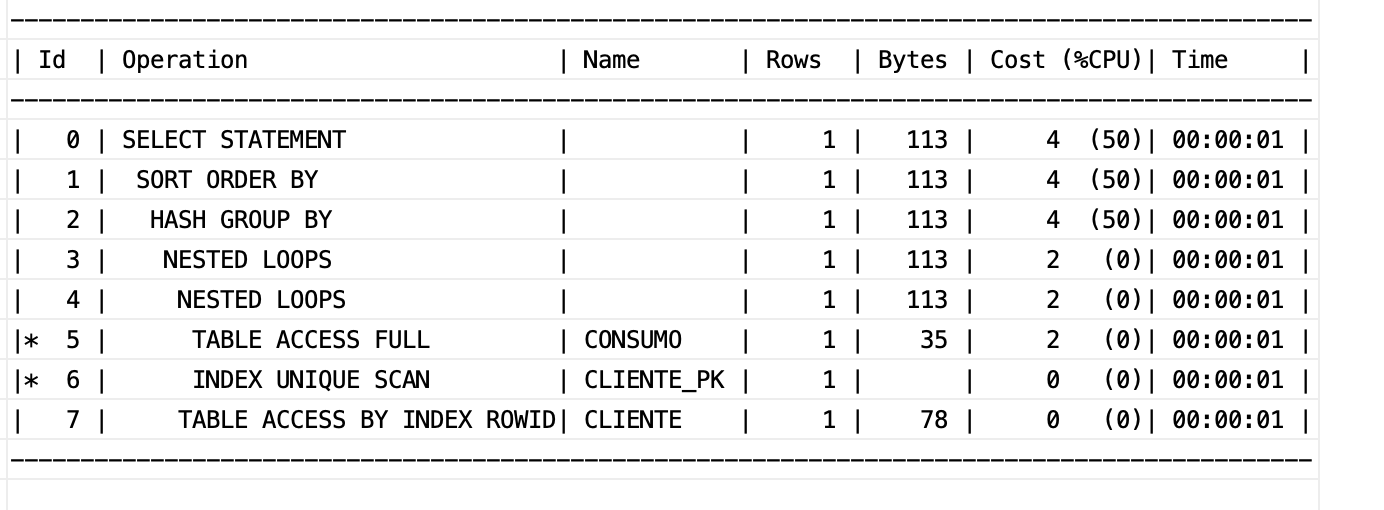
GROUP BY c.ID, c.DIRECCION, c.TELEFONO

ORDER BY COUNT(\*) DESC;

Plan de ejecución



Tiempo de ejecución:



**RFC10 - CONSULTAR CONSUMO EN HOTELANDES – RFC9-V2**

SELECT c.ID AS CLIENTE\_ID, c.DIRECCION, c.TELEFONO, COUNT(\*)

FROM cliente c

LEFT JOIN consumo co ON c.ID = co.CLIENTE\_ID

AND co.SERVICIO\_ID = 5

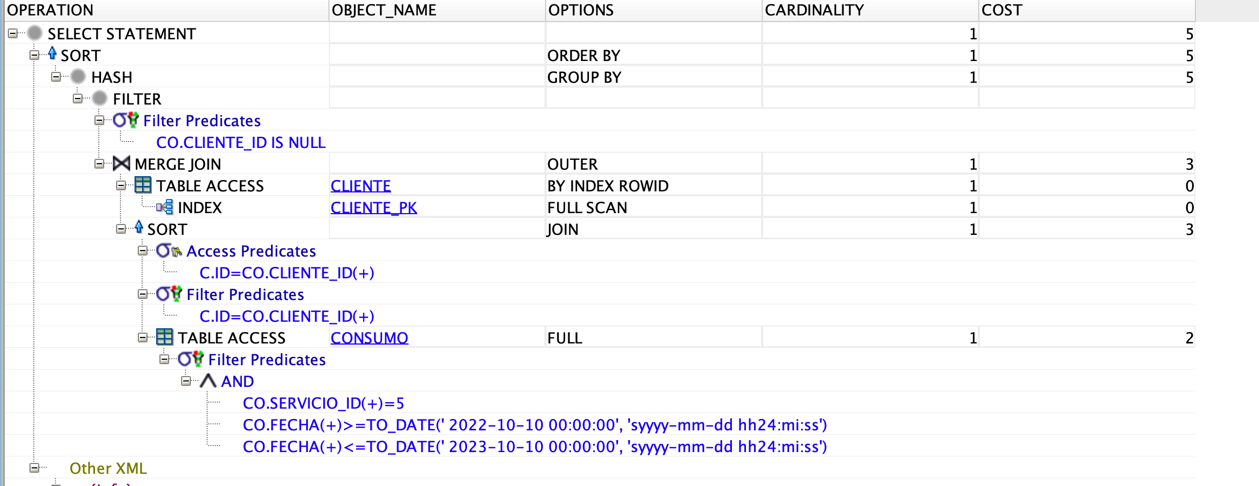
AND co.FECHA BETWEEN '10/10/2022' AND '10/10/2023'

WHERE co.CLIENTE\_ID IS NULL

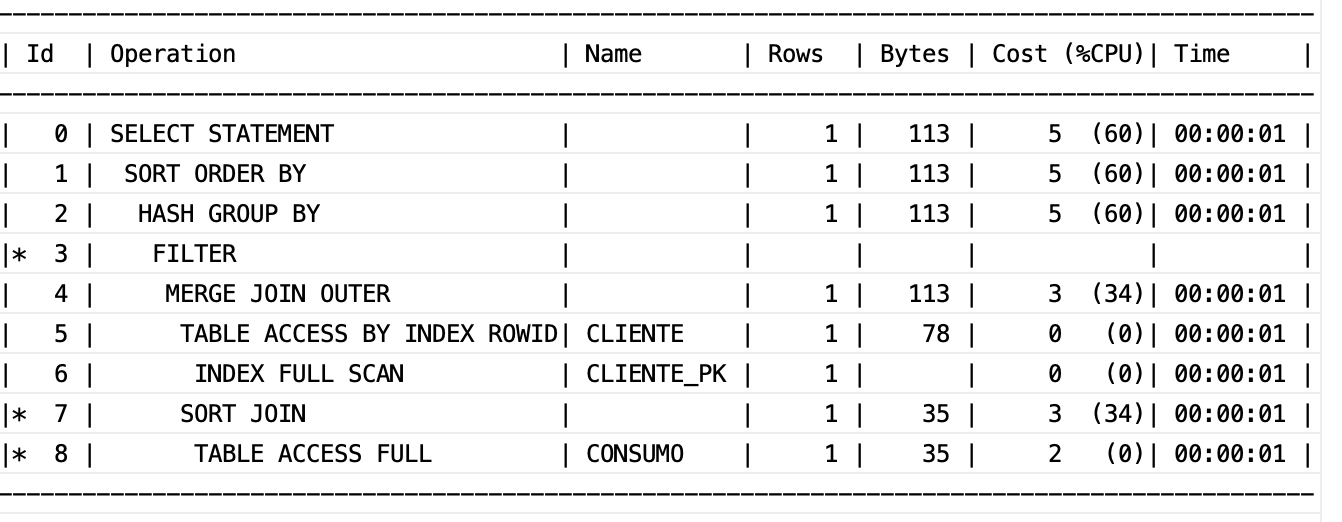
GROUP BY c.ID, c.DIRECCION, c.TELEFONO

ORDER BY COUNT(\*) ;

Plan de ejecución:



Tiempo de ejecución:



**RFC12 - CONSULTAR LOS CLIENTES EXCELENTES**

SELECT cliente\_id

FROM ((

SELECT r.cliente\_id

FROM reserva r

GROUP BY r.cliente\_id, EXTRACT(YEAR FROM r.fechaentrada)

HAVING COUNT(DISTINCT EXTRACT(YEAR FROM r.fechaentrada)) > 4

)

UNION

SELECT c.ID AS cliente\_id

FROM cliente c

WHERE EXISTS (

SELECT 1

FROM consumo co

WHERE co.CLIENTE\_ID = c.ID

AND co.PRECIOUNITARIO > 300000

)

UNION

SELECT r.cliente\_id

FROM reserva r

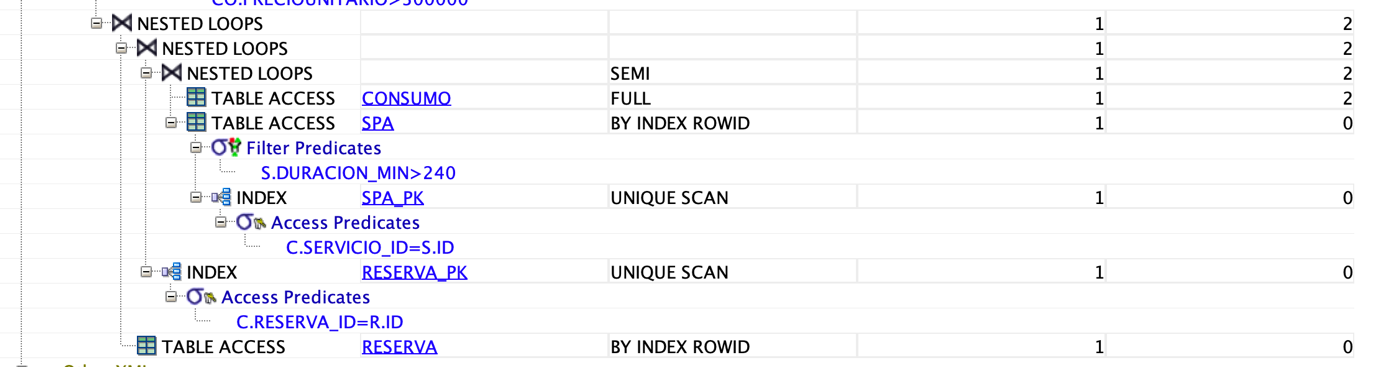
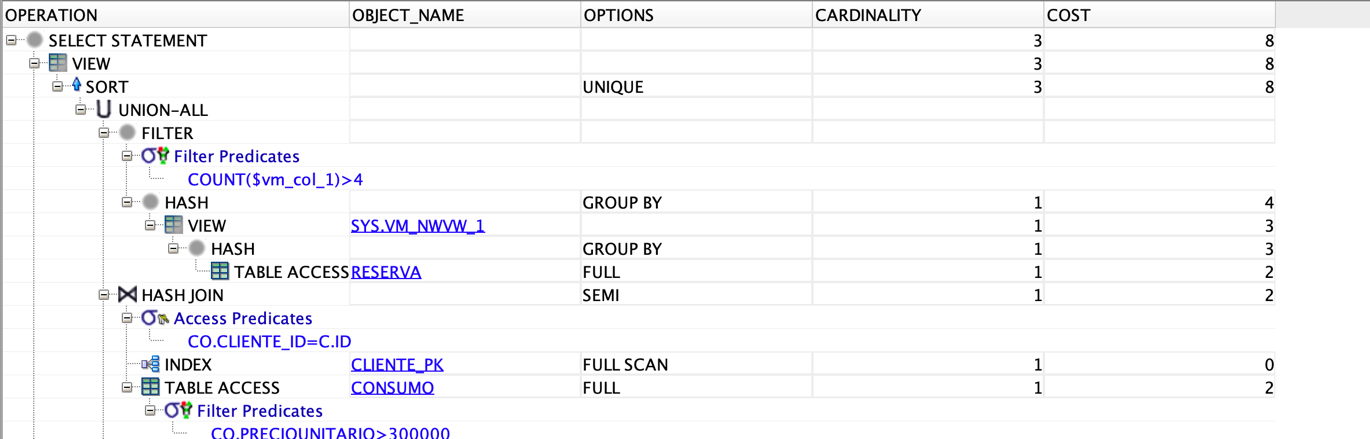
JOIN consumo c ON c.RESERVA\_ID = r.ID

JOIN spa s ON c.SERVICIO\_ID = s.ID

WHERE s.DURACION\_MIN > (4 \* 60)

);

Plan de ejecución:



Tiempo de ejecución:

