Diseño de la aplicación:

1. Para esta iteración el modelo planteado desde la segunda iteración no ha presentado cambios drásticos en cuanto a estructura relacional y al respaldo de persistencia de la aplicación. Los nuevos requerimientos funcionales a realizar eran únicamente de consulta por lo que no era necesario implementar modificaciones en la BD. Por el otro lado, para la iteración si fue necesario la creación de índices sobre algunas columnas de las tablas con el fin de volver el tiempo y costo de acceso más eficiente debido a la cantidad de datos utilizados.

Diseño físico:

Justificaciones índices.

o RF9

En el noveno requerimiento era necesario hacer una consulta donde se involucraba nuestras tablas: compania, usuarios, funciones, boletas y espectaculos. Al incluir estas 5 tablas y cada una poblada con al menos 10000 datos era necesario imprentar índices sobre las columnas por la cuales se filtraba algún parámetro y aquellas donde se realizaba el inner o natural Join. Fue de esta manera que decidimos implementar índices B+, bitmap, primarios y compuestos. En primer lugar, utilizamos índices B+, o conocidos en Oracle Sql Developer como "normal", para aquellas columnas que son llaves primarias. Las implementaciones de índices para llaves primarias son de gran ayuda sobre todo en las búsquedas sobre predicados de igualdad (=) debido a que el IndexScan que se realiza es único, pues tanto la cardinalidad como el costo es bajo. De igual forma, tanto el costo como la cardinalidad se reducen drásticamente utilizando un índice normal pero compuesto. Estos fueron usados en la tabla espectaculo sobre las columnas id_Espectaculo e ID_Compania, se da prelación al id_espectaciulo para filtrar por predicado con el fin de realizar el hash join, y el ID_espectaculo para cumplir con el predicado de igualdad, mediante una búsqueda de rango con el índice. En cuanto a índices normales simples se encuentra el índice aplicado sobre la tabla función y la columna id_espectaculo con el fin de realizar una igualdad de predicado más eficaz. Por último se implementaron índices Bitmap para los campos ID_Funcion e ID_Silla en la tabla boleta donde vale destacar que la única llave primaria es ID_Boleta, por ende al realizar un índice bitmap sobre estos campos realiza un escaneo de rango mucho más eficiente en espacio y tiempo.

() INDEX_OWNER	() INDEX_NAME	() UNIQUEN	ISS () STAT	TUS () INDEX_1	TYPE () TEMPOR	RARY () PARTITIO	MED @ FUNCTOX_STAT	US () JOIN_ING	EX () COLUMNS	COLUMN_EXPRESSI
1 ISIS2384A241728	ID_ESPEC_PK	UNIQUE	VALID	MORMAL	N	NO	(null)	140	ID_ESPEC	(null)
2 ISIS2384A241728	INDEX_ESPEC_CO	IPA NONUNIQUE	VALID	MORMAL	N	NO	(null)	NO	ID_ESPEC, ID_COMPAR	EIA (null)
() INDEX.OWNER	() INDEX NAME	UNIQUENESS	STATUS.	() INDEX TYPE	A TEMPORAR	v ili pastimosen	0 FUNCION STATUS	() ION INDE	A COLUMNS	COLUMN EXPRESSION
1 ISIS2384A241728				BITHAP	N	NO	(null)	NO	ID_FUNCTION, ID_STILLA	
2 ISIS2384A241728				NORMUL.	N	NO	(null)	NO	ID_BOLETA	(null)
3 ISIS2384A241728	ESTADO_DVDEX	NONUNDQUE	WLID	BITHAP	N	NO	(null)	NO	ESTADO	(null)
4 ISIS2384A221728	INDEX_IDSILLA	NONUNDQUE	WLID	NORMAL	M	NO	(null)	NO	ID_SILLA	(null)
() INDEX_OWNER	INDEX_NAME	() UNIQUENESS) STATUS	() INDEX_TYPE) TEMPORARY	() PARTITIONED ()	FUNCION_STATUS () JO	INJINDEX () O	OLUMNS	COLUMN_EXPRESSION
1 15152384A241728 1	NOEX_TOESPEC	NONUNTOUS	WALID	NORMAL.	N	ND (r	nutt) NO	10_	ESPECTACULO	(null)
2 15152384A241728 1	IO_FUNCTION_PK	UNITQUE	WLID	NORMAL.	N I	NO 0	null) NO	10_	FUNCTION	(null)
3 ISIS2384A221728 I	MORY RIM PERSON	MONE INTO UE	WALID	NORMAL.	N I	ND D	utt) NO	10	PUNCTON, ID ESPECTACULO	(end1)

o RF10

Debido a nuestro inicial diseño respecto a la BD, este requerimiento se lograba cumplir negando una sentencia del RF9. Esta sentencia es aquella donde se dice que el estado ='A' donde 'A' significa activo. Para este requerimiento el único cambio necesario fue estado='D' donde 'D' significa devuelto, es decir que el usuario no asistió a la función. Por ende, la explicación de índices y el plan de consulta es básicamente el mismo que el RF9.

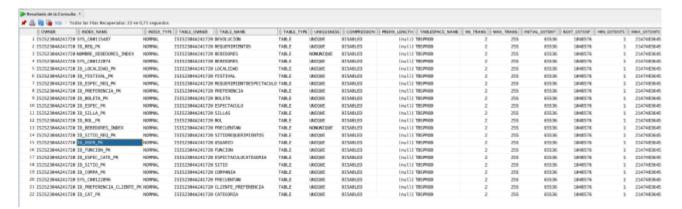
o RF11

El requerimiento 11 debía mostrar la información de las boletas dependiendo de ciertos críterios dados por el usuario sobre la función. Para ello se decidío implementar índices sobre los críterios de busqueda y sobre aquellos atributos o columnas donde se realizan los joins. En cuanto a los índices implementados y utilizados para esta sentencia, se utilizó un índice compuesto de tipo normal o B+ para la tabla requerimiento Espectaculo aplicado a las columnas ID_requerimiento y ID_Espectaculo. Debido a que en la sentencia se realiza un inner join entre la tabla 5 y requerimientos donde la columna de join es id_requerimiento, la tabla 5 es el resultado de un inner join entre requerimiento Espectaculo y la tabla 4 donde la columna igualdad es esl id_espec. Por tal motivo se decidió implementar tales índices en ese orden. En la mitad de la sentencia se ejecuta un inner join entre función y espectáculo mediante el id_espec, luego se realiza un acceso a predicados en un intervalo dado. El índice aplicado para esta parte de la sentencia causa en la disminución drástica de costos de ejecución, esto se da por realizar el id compuesto entre ID_Espectaculo y hora.

() INDEX_OWNER	() INDEX_NAME	i) uniqu	DVESS () ST	TATUS ()	INDEX_TYPE	TEMPORARY	PARTITIONED	() FUNCION_STATUS	() JOHN_INDEX	() COLUMNS	COLUMN_DIPRESSIO
1 15152304A241720	DIDEX_IDESPEC	MONUNGO	UE VAL	30 N	DRIVAL 1	M	MO	(mull)	NO NO	ID_ESPECTACULO	(mull)
7 15152384A241728	ID_PUNCTON_PK	UNIQUE	VAL	ID N	DRMAL I	M	MO	(mult)	140	ID_FUNCION	(mult)
3 15152384A241728	INDEX_FUN_ESPEC	HORA MONUNCO	UE VAL	30 N	DRIVAL 1	M	MO	(mult)	NO CO	ID_ESPECTACULD, HDRA	(mull)
4 15152304A221720	DIDEX_FUNC_ESPE	C MOMUNES	UE VAL	30 N	DRIVAL	м	MO	(mull)	140	ID_FUNCTION, ID_ESPECTACU	LO (mull)
4 15352384A221728 @INDEX_OWNER	DIOEX_FUNC_ESPE							(mull) XCSTATUS () JOHN IN			COLUMN_DOPRESSION
	() INDEX_MAKE	UNIQUENESS	STATUS						DEX () COLUM		COLUMN_EXPRESSION
	() INDEX_MAKE	UNIQUENESS	STATUS (NOEX_T	TYPE () TEMPO		DONED () FUNCE	X_STATUS () JOIN, IN	DEX () COLUM ID_PEQUE	NS)	COLUMN_EXPRESSION

o RF12

Índices generados automáticamente por Oracle



Cuando se crea una tabla Oracle automáticamente genera un índice único sobre la columna de la llave primaria para generar el constraint de que no se puedan repetir. Como podemos ver en la imagen hasta el momento solo existen indices para la llave primaria de cada tabla. Estos índices si mejoran el rendimiento de las sentencias ya que por ejemplo en las sentencias 1,2 y 4 debemos dar la información de los clientes que tienen ciertas características sobre sus boletas por ende después de hacer el filtro que nos permite conocer cuáles son estos clientes debemos buscarlos en la tabla de usuarios y estos índices nos ayudan a que no tengamos que buscar en toda la tabla sino llegar solamente a los índices que coincidan con lo que estamos buscando.

Documentación escenarios de prueba

Requerimiento 9:

• Sentencia: with tabla1 as (select id_funcion as id_fun from ((select id_espec from ISIS2304A241720.espectaculo where id_compania = 19)

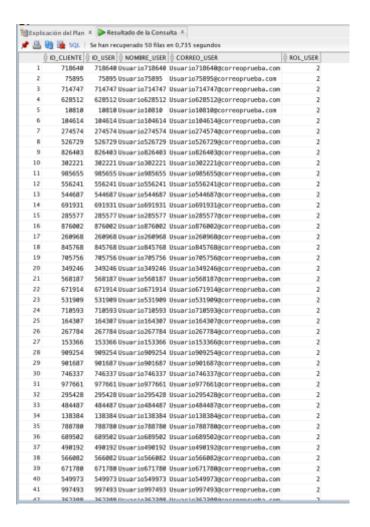
inner join ISIS2304A241720.funcion on id_espec = id_espectaculo)), tabla2 as (select distinct id_cliente from tabla1 inner join (select * from ISIS2304A241720.boleta where estado = 'A') on id_fun = id_funcion)

select * from (tabla2 inner join ISIS2304A241720.usuario on id_cliente = id_user);

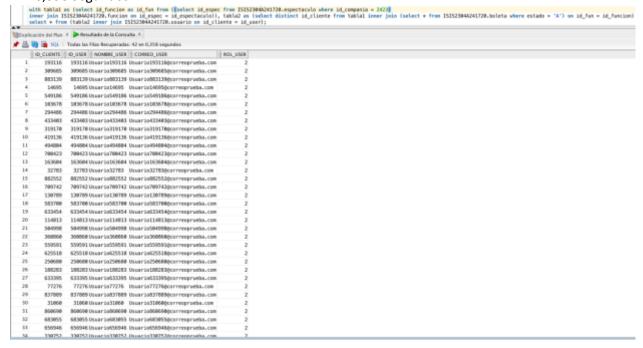
- Distribución de los datos: Los datos estan distribuidos asimetricamente por lo cual las sentencias no tienen la misma duración. Esto quiere decir que no hay el mismo numero de boletas para todas las funciones, y que no todas tienen boletas tampoco. Por ende, dependiendo del numero de la compañía que entre por parametros los tiempos de ejecucion no seran los mismos gracias a los indices. Ademas tampoco es simetrica entre las boletas asistidas y las devueltas por lo cual esto tambien interfiere en el tiempo de ejecucion de la sentencia.
- Valores de los parámetros: id compania = 8622 y 2423 los tiempos de ejecucion son distintos, aunque el plan es el mismo debido a la asimetria de los datos



• Tiempo de ejecución: t1=0,735 segundos



T2= 0,358 Segundos



Requerimiento 10

• Sentencia: with tabla1 as (select id_funcion as id_fun from ((select id_espec from ISIS2304A241720.espectaculo where id_compania = 19)

inner join ISIS2304A241720.funcion on id_espec = id_espectaculo)),

tabla2 as (select distinct id_cliente from tabla1 inner join (select * from ISIS2304A241720.boleta where estado = 'D') on id fun = id funcion)

select * from (tabla2 inner join ISIS2304A241720.usuario on id_cliente = id_user);

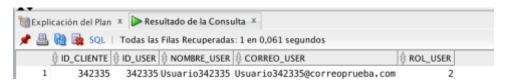
- Distribución de los datos: Así como en el literal anterior los datos no estan distribuidos simetricamente en cuanto a la cantidad de boletas que tiene cada compañía por cliente esto quiere decir que dependiendo de la compania que busquemos los tiempos de ejecucion seran distintos debido a los indices que nos permiten llegar a esta informacion
- Valores de los parámetros: Para esta usamos id compania= 4564 y 1131



• Tiempo de ejecución: t1=0,053 segundos



T2 = 0,061 segundos



Requerimiento 11

- Sentencia: with tabla1 as (select * from sillas natural join localidad), tabla2 as (select * from boleta natural join funcion), tabla3 as (select *from tabla1 inner join tabla2 on tabla1.id_silla = tabla2.id_silla), tabla4 as (select distinct * from tabla3 inner join espectaculo on tabla3.id_espectaculo = espectaculo.id_espec), tabla5 as (select * from tabla4 inner join requerimientoespectaculo on tabla4.id_espec = requerimientoespectaculo.ID_ESPECTACULO) select * from tabla5 inner join REQUERIMIENTOS on REQUERIMIENTOS.ID_REQ = tabla5.id_requerimiento where fecha between '02/02/2017' and '05/07/2017';
- Distribucion de los datos: Para este requerimiento tenemos muchos mas criterios por los cuales evaluar la sentencia dependiendo de las preferencias del cliente. Sin embargo la tabla no esta simetricamente distribuida por estos parametros. Debido a esto las sentencias no se demoraran lo mismo porque la cantidad de datos que encontrara gracias a los indices no siempre sera la misma
- Valores de los parámetros: 1350>Franja horaria >1120

Plan de ejecución: CARDINALITY SELECT STATEMENT 11191 HASH JOIN

G. Access Predicates
SELAS JD_LOCALIDAD SELASID, JOCALIDAD-LOCALIDADID, LOCALIDA

TARE ACCESS OFULL

HASH, JON

SON, ACCESS THE SELASID, SELA

TABLE ACCESS OF ULL

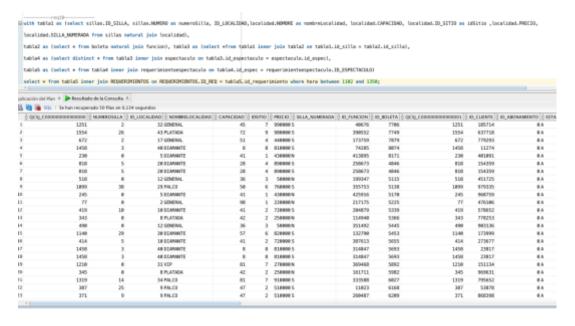
OT ACCESS THE SELASID, SELA

OT ACCESS THE SELASID, SELASID 7224 1813 - On Access Predicat COSTS PredICARIES
FUNCION.ID_ESPECTACULO=ESPECTACULO.ID_ESPEC 7224 1813 M DE HASH JOIN 1018 1093 1018 1024 (Tip. Access Predicates NESTED LOOPS

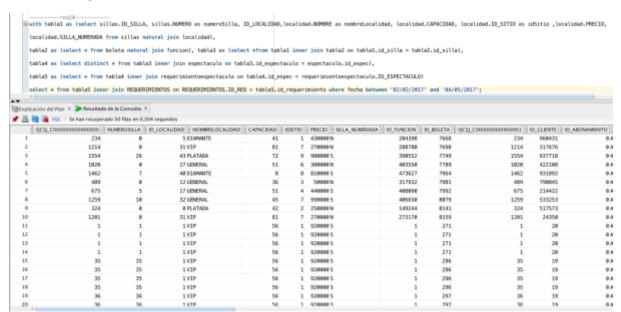
B M NESTED LOOPS

B STATISTICS all INDEX (FAST FULL SCAN)
all INDEX (UNIQUE SCAN)
all Obs. Access Predicates REQUERIMIENTOS ESPECTACULO TABLE ACCESS (FULL) 11020 On Access Predicates FUNCTION ID ESPECTACULO «ESPECTACULO JO ESPEC FUNCION HORAS = 1102 FUNCION HORA <= 1350 TABLE ACCESS (FULL) FUNCION 78170 TABLE ACCESS (FULL) BOLETA 498958

Tiempo de ejecución: t1=0,134 segundos

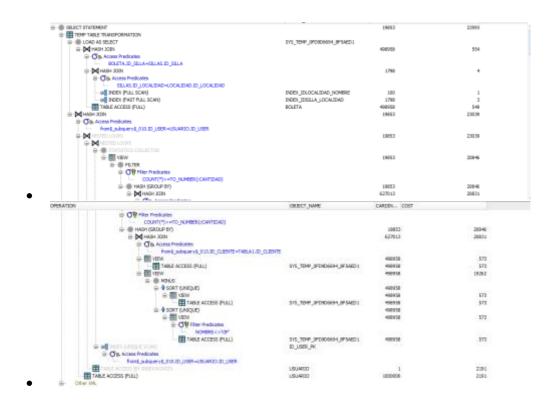


T2 = 0,304 segundos



Requerimiento 12

- Sentencia: with tabla1 as (select id_boleta, id_cliente, nombre from (select id_boleta, id_silla, id_cliente, id_localidad from ((select * from ISIS2304A241720.boleta) natural join (select * from ISIS2304A241720.localidad), tabla3 as(select id_boleta, id_cliente, nombre from tabla1 where nombre != 'VIP') select * from ((select id_cliente as id_user from(select distinct id_cliente, count(*) as cantidad from(Select * from (select id_cliente from tabla1 minus select id_cliente from tabla3) natural join tabla1) group by id_cliente) where cantidad >= &cantidad) natural join ISIS2304A241720.usuario);
- Distribución de los datos: Los datos de las boletas no estan simetricamente distribuidos con respecto a las localidades de las boletas ni porque un cliente compre solo un tipo de localidad por esto las diferentes cantidades que pongamos nos traeran diferentes resultados y por ende tomaran tiempos distintos
- Valores de los parámetros: cantidad = 4, cantidad = 6
- Plan de ejecución



CASOS DE PRUEBA:

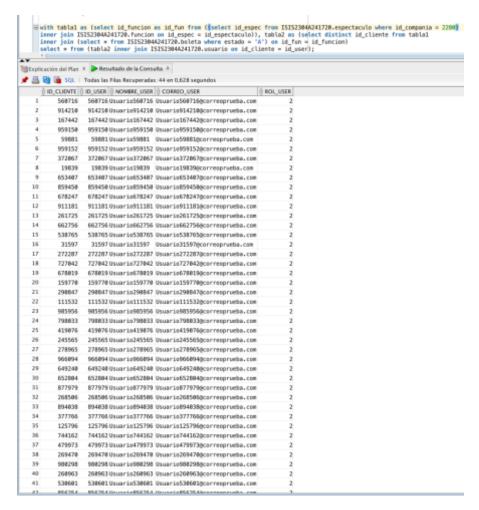
1. RF9:

a. Caso de prueba 1:

Se realizará el requerimiento funcional de consulta número 9 con el parámetro de consulta tal que el id de la compañía es el 2280. Se espera como resultado al menos una tupla con la información requerida por el requerimiento.

El plan de ejecucción que creemos que sucederá es: Debido a que la sentencia tiene varias vistas de tablas (views) involucradas, se resolverán primero los select * más internos de la sentencia donde estos serían select * from espectaculo where id_compania = 2280, para esta sentencia se espera que se haga el filtro de predicado usando el índice index_espec_compa. Luego se espera que resuelva un inner join con la tabla función bajo los valores de id_espectaculo donde se pretende que use el mismo índice anteriormente descrito. Después, se espera que para realizar la sentencia que contiene la igualdad estado='A' use el índice bitmap creado sobre la columna estado para resolver la sentencia mucho más eficaz. Con el resultado de la anterior sentencia, puede ocurrir el inner join con la tabla cliente bajo el id_cliente como relación de las dos tablas. Valga aclarar que se utiliza una columna Estado para identificar si la boleta aún está activa o fue devuelta antes que se realizará la función

Resultado: 44 tuplas recuperadas en 0,620 segundos



Plan de ejecución:



De acuerdo al plan de ejecución realizado por Oracle se tiene que: en primera medida si utilizó el índice index_espec_compa para filtrar por el predicado 2280. Luego de esto, utiliza el índice creado llamado index_idEspec el cual es de tipo normal. Línea despues utiliza un índice sobre la tabla boleta el cual es un índice compuesto donde se resuelve una parte del nested loop. Por úlitmo no utiliza el índice del predicado bitmap sobre la columna estado para realizar la respectiva comparación.

2. RF10:

a. Caso de prueba 1:

Se realizará el requerimiento funcional de consulta número 10 con el parámetro de consulta tal que el id de la compañía es el 2116. Se espera como resultado al menos una tupla con la información requerida por el requerimiento.

Para la resolución de esta sentencia se espera lo mismo que el anterior requerimiento funcional con la diferencia de que cambia el valor del estado en la sentencia: Debido a que la sentencia tiene varias vistas de tablas (views) involucradas, se resolverán primero los select * más internos de la sentencia donde estos serían select * from espectaculo where id_compania = 2280, para esta sentencia se espera que se haga el filtro de predicado usando el índice index_espec_compa. Luego se espera que resuelva un inner join con la tabla función bajo los valores de id_espectaculo donde se pretende que use el mismo índice anteriormente descrito. Después, se espera que para realizar la sentencia que contiene la igualdad estado='D' use el índice bitmap creado sobre la columna estado para resolver la sentencia mucho más eficaz. Con el resultado de la anterior sentencia, puede ocurrir el inner join con la tabla cliente bajo el id_cliente como relación de las dos tablas. Valga aclarar que se utiliza una columna Estado para identificar si la boleta aún está activa o fue devuelta antes que se realizará la función.

Resultado: 1 tupla en 0,089 segundos.



De acuerdo al plan de ejecución realizado por Oracle se tiene que: en primera medida si utilizó el índice index_espec_compa para filtrar por el predicado 2116. Luego de esto, utiliza el índice creado llamado index_idEspec el cual es de tipo normal. Líneas después utiliza un índice sobre la tabla boleta el cual es un índice compuesto donde se resuelve una parte del nested loop. Por úlitmo y a diferencia del anterior requerimiento,

se utiliza el índice bitmap sobre la columna estado para encontrar todas aquellas tuplas en la cual el estado es "D".

3. RF11

a. Caso de prueba 1.

Se realizará el requerimiento número 11 haciendo un filtrado de datos por la franja horaria. Estos datos serán entre las 20 horas y las 24 horas. Se espera como resultado al menos una tupla con la información solicitada en el requerimiento.

El plan de ejecución esperado para el requerimiento funcional 11 al igual que los otros requerimientos, se inicia desarrollando las sentencias internas haciendo especial énfasis en los natural e inner joins donde se espera que utilice el índice compuesto creado para la tabla requerimientoEspectaculo. De igual forma, se espera que al momento de ejecutar la sentencia where para la columna de franja horaria, se utilice el índice compuesto creado especialmente para ese requerimiento, tal índice está formado por la hora y el id del espectáculo donde se le da prevalencia al campo de hora.



b. Caso de prueba 2:

Se realizará el requerimiento número 11 haciendo un filtrado de datos por la Fecha. Estos datos serán entre la fecha 02/05/2017 y las 02/12/2017. Se espera como resultado al menos una tupla con la información solicitada en el requerimiento.

Para el plan de ejecución al igual que el caso de prueba 1, se espera que se inicia desarrollando las sentencias internas haciendo especial énfasis en los natural e inner joins donde se espera que utilice el índice compuesto creado para la tabla requerimientoEspectaculo. Al momento de ejecutar la sentencia del where con datos de

fechas, se espera que primero se organice la tabla mediante esta columna y luego realice la búsqueda dentro del rango establecido.

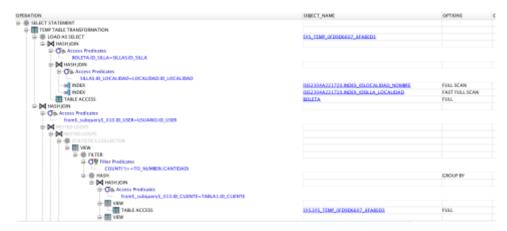


4. RF12

a. Caso de prueba 1:

Se realizará el requerimiento número 12 de acuerdo al enunciado, tomando las boletas de los clientes las cuales no son VIP y luego "restando" este resultado de la tabla total de boletas con el fin de dejar únicamente los clientes que si compraron boleta en localidad "VIP".

El de ejecución esperado para la sentencia es: primero evaluar las sentencias donde sea necesario realizar hash join o nested loops mediante llaves primarias con índices previamente creados. Luego se espera que se evalué mediante los índices creados en la tabla localidad con las columnas id localidad y nombre con el fin de efectuar de una manera más eficaz las sentencias join mediante el id y los condicionales where con el nombre de la localidad. De igual forma se espera la ejecución de la sentencia usando el índice creado sobre la tabla silla para las columnas id silla e id localidad con el de fin de que cumplan la misma función descrita anteriormente.





POBLAMIENTO BD:

En un principio, tocó generar los datos de manera que en las tablas hubiera información coherente. Por ello, creamos un pequeño programa generador de datos. En este nosotros podíamos decidir la cantidad de datos que generaban, y en qué formato quedaba. Decidimos escribir dentro de archivos .csv debido a que nos pareció la mejor opción.

Se crearon alrededor de un millón de datos para las tablas: Boleta, Boleta_silla, Boleta_detalle y Usuarios. Las demás tablas se les generaron datos de tal manera que el manejo de la coherencia de los datos fuera fácil y rápido.

Teniendo los datos ya creados, nos dirigimos a la plataforma SQL developer donde se encuentran nuestras tablas del proyecto. Para cada tabla a la que se le generaron datos, se produjo el proceso de acceder a la opción de importación de datos de la tabla. En esta seleccionamos la ruta en donde se encuentra el archivo .csv que generó el programa anteriormente. Luego de seguir una serie de pasos dados por el Wizard de SQLDeveloper. La importanción de alrededor un millón de datos por tabla fue efectiva.

Análisis resultados:

Si trajéramos todos los datos a memoria principal realizar todas las operaciones de filtro y joins sería mucho menos eficiente y el programa ocuparía mucha más memoria. Primero deberíamos traer toda la información de las tablas y guardarlos objetos y estos en estructuras como arraylist o hash tables donde usaríamos los id como key para estas tablas. Cuando ya hayamos traído toda la información podríamos comenzar a filtrar por los criterios que necesitáramos pero tendríamos que usar estructuras temporales para guardar esta información lo cual ocuparía mucha más memoria. Además de que no podríamos filtrar y unir las tablas tan fácilmente porque primero necesitaríamos filtrar una, luego la otra y luego intentar unirlas y guardarlas en nuevos objetos que nos permitan guardar la información que estamos uniendo. Esto requeriría mucho as espacio y tiempo de ejecución cuando estamos hablando de una cantidad de datos considerable.