

# Uso de índices

## Introducción

El objetivo del laboratorio es probar la incidencia del uso de los índices en el tiempo de respuesta de las consultas/actualizaciones sobre “tablas grandes”.

Para ello, se ha elegido un dominio consistente en dos tablas, **ALUMNOS** y **MATRICULACIONES**. La primera contiene el DNI, nombre, apellidos y fecha de nacimiento de los alumnos; mientras que la segunda almacena el DNI del alumno, la referencia numérica de la asignatura en la que se ha matriculado, y el curso académico en el que lo ha hecho (Se ha usado una clave artificial ID para evitar trabajar con una clave compuesta). En concreto, esta es la definición SQL de las tablas que se utilizarán:

```
CREATE TABLE "ALUMNOS"
(
    "DNI" NUMBER NOT NULL,
    "NOMBRE" VARCHAR2(13),
    "APELLIDOS" VARCHAR2(30),
    "NACIMIENTO" DATE,
    PRIMARY KEY ("DNI")
);

CREATE TABLE "MATRICULACIONES"
(
    "ID" NUMBER(8) NOT NULL,
    "DNI" NUMBER(8),
    "ASIGNATURA" NUMBER,
    "CURSO" NUMBER,
    PRIMARY KEY ("ID")
);
```

NOTA: EL archivo example.bat se encuentra en:

<https://bit.ly/difrepositorio>

password: difbiltegia

SOFTWARE/software2020-2021/ORACLE12C/jdbc\_oracle

Bajar al escritorio la carpeta completa jdbc\_Oracle

## Tareas

- 1.- Hacer doble click sobre el archivo *example.bat* para lanzar el programa.
- 2.- Introducir el nombre de **usuario y contraseña** que corresponda.

### Nota

Los nombres de usuario (y las contraseñas) de **Oracle** deben escribirse en **mayúsculas**.

- 3.- En la lista de opciones, seleccionar **Crear las tablas ALUMNOS y MATRICULACIONES** (opción 1). Así se crearán automáticamente las dos tablas mencionadas anteriormente.
- 4.- Ahora seleccionar **Insertar tuplas en la tabla ALUMNOS** (opción 2). El programa preguntará el número de tuplas que se desean insertar en la tabla. Seleccionar **2000000** (tiempo aproximado de espera estimado: 1minuto). Cuando la operación se complete, volverá a aparecer el mismo diálogo.
- 5.- Seleccionar la opción 3 (**Insertar tuplas en la tabla MATRICULACIONES**) para llenar también la tabla **MATRICULACIONES**. Seleccionar la misma cantidad que antes: **2000000** tuplas. Tardará un tiempo similar al caso anterior.
- 6.- Para conocer el tiempo que consumen las consultas/actualizaciones  
Entrar en SQL Worksheet y poner SET TIMING ON AUTOTRACE ON y a partir de aquí al ejecutar las preguntas nos saldrá información referente a la misma.  
Para eliminar esta opción SET AUTOTRACE OFF
- 7.- Ejecutar desde SQL Worksheet la consulta que permita contar el número de alumnos con nombre = 'Alfredo'

Ejecutar la consulta anterior varias veces más (por ejemplo, 4 veces) y observar el tiempo tardado en cada una de ellas. **Anotar el resultado**. En caso de que una de las ejecuciones (concretamente, la primera) haya tardado significativamente más que el resto, anotar también ese valor. ¿A qué puede deberse?

Oracle cachea los resultados de las preguntas. Para eliminar de la cache ejecutar la sentencia

```
ALTER SYSTEM FLUSH BUFFER_CACHE;
```

Volver a ejecutar la pregunta para ver la variación en el tiempo de respuesta.

- 8.- Escribir y ejecutar la consulta que permita contar el número de alumnos con DNI= '10999999' (Los valores del DNI están entre 10000000 y 11999999);

Ejecutar la consulta anterior varias veces más (por ejemplo, 4 veces) y observar el tiempo tardado en cada una de ellas. **Anotar el resultado**. En caso de que una de las

ejecuciones (concretamente, la primera) haya tardado significativamente más que el resto, anotar también ese valor. ¿A qué puede deberse?

9. ¿Se observa alguna diferencia entre los tiempos en las preguntas de los apartados 7 y 8?

10.- Crear un índice sobre el campo **NOMBRE** ejecutando la siguiente sentencia SQL:

```
CREATE INDEX NOMBRES ON ALUMNOS (NOMBRE)
```

Anotar el tiempo consumido.

¿De qué tipo de índice se trata?

11.- Borrar la cache. Probar varias veces la consulta ejecutada en el paso 7. Anotar el resultado y compararlo con el resultado del apartado 7. ¿Se nota el uso del índice?

12.- La tabla **ALUMNOS** tiene como clave primaria el campo **DNI**. Al definirlo como tal, el SGBD crea automáticamente un índice primario para dicho campo, con el objetivo de por un lado preservar la restricción de unicidad, y por otro acelerar los **JOIN** entre tablas en las que participe la clave primaria.

Suponiendo que se desconoce esto, intentar crear un índice para el campo **DNI**, tal y como se ha hecho anteriormente con el campo **NOMBRE**.

### **Anotar lo que ocurre**

13.- Para comprobar la influencia de los índices en los **JOIN**, (primero borrar cache)

a) ejecutar la siguiente consulta: Obtener el nombre de los alumnos que, siendo su DNI entre 10000000 y 11999999, se hayan matriculado hasta ahora de más de 10 asignaturas.

Ejecutar varias veces la consulta y **anotar el resultado**.

b) Borrar cache y definir un índice sobre el campo DNI de Matriculaciones

14. Borrar cache y borrar el índice sobre el campo Nombre de la tabla **ALUMNOS** y crearlo de tipo BitMap

```
CREATE BITMAP INDEX NOMBRES ON ALUMNOS (NOMBRE)
```

```
DROP INDEX NOMBRES
```

En un índice bitmap, en vez de una lista de identificadores de tupla, Oracle crea un mapa de bits para cada valor clave del índice. Recomendable para atributos de baja cardinalidad (ejemplo sexo, estado civil). Ocupan menos espacio.

Ejecutar la pregunta del apartado 7

¿El tiempo de procesamiento de la pregunta ha mejorado?

15. Borrar cache. Borrar el índice creado en el apartado anterior y crear un índice sobre el campo nombre de tipo Reverse.

Reverse Index. Toma los valores del atributo y les da la vuelta antes de introducirlos en el índice. Ej. ID 123456 Oracle lo transforma en 654321 antes de introducirlo al índice

Parece que mejora la distribución de los valores en el espacio global.

```
CREATE INDEX Nombre ON ALUMNOS (NOMBRE) REVERSE
```

Repetir la ejecución de la pregunta del apartado 7. Ver los tiempos

16. Para ver los índices creados sobre la tabla ALUMNOS

```
select index_name, index_type, table_name, dropped  
from all_indexes  
where table_name = 'ALUMNOS';
```

17. Borrar cache y borrar los índices creados por vosotros en la tabla Alumnos

Ejecutar la pregunta que cambia la fecha de nacimiento del alumno de nombre Alfredo

```
UPDATE ALUMNOS SET Nacimiento = '28-MAR-95', Apellidos = 'Martinez' where  
Nombre = 'Alfredo';
```

Crear dos índices para la tabla ALUMNOS, uno para Nacimiento y otro para Apellidos;

```
CREATE INDEX Nac ON ALUMNOS (NACIMIENTO)
```

```
CREATE INDEX Apes ON ALUMNOS (APELLIDOS)
```

Borrar cache y ejecutar la pregunta anterior y comparar los tiempos obtenidos. ¿Han mejorado los tiempos?

Borrar las tablas: ALUMNOS y MATRICULACIONES.