

Administración de Bases de Datos – Traballos tutelados

TT2: Axuste da base de datos

Curso 2023/2024

Data límite para a entrega: 24/05/2024

1. Contextualización

O obxectivo xeral dos traballos tutelados en ABD é deseñar unha pequena base de datos, que será posteriormente implementada utilizando Oracle, alimentada con datos, e afinada para que se comporte de xeito eficiente.

Para iso realizaranse 2 traballos:

TT1: (1,2 puntos) Deseño, implementación e alimentación dunha base de datos.

Realizarase o deseño conceptual e lóxico da base de datos. Posteriormente a base de datos será implementada en Oracle, e alimentada con datos a través de sentencias SQL e con ferramentas externas como SQL*Loader.

TT2: (1,8 puntos) Axuste da base de datos.

Debe describirse unha carga de traballo habitual para a base de datos, indicando unha serie de consultas frecuentes que quereremos optimizar.

Utilizando esa información, a base de datos debe ser afinada para que responda coa maior eficiencia posible a esa carga de traballo.

Como tamaño orientativo para a base de datos, esta debería conter ó redor de 5 táboas unha vez transformada ó modelo relacional.

2. TT2: Traballo a realizar

O traballo a realizar para esta segunda entrega consiste no axuste e optimización da base de datos creada no TT1, en base a unha carga de traballo que hai que definir. Esta carga de traballo consistirá en 5 sentencias `select` que sexan frecuentes para esa base de datos, máis unha sentencia DML que utilizaremos para unha verificación final do axuste realizado.

En concreto, o traballo a realizar será o seguinte:

1. Para cada unha das 5 sentencias `select` que definen a carga de traballo:

- a) Especificar o enunciado da consulta. En teoría, sería unha consulta frecuente para a BD definida pero, dado que é unha base de datos ficticia, a carga de traballo será estimada. En calquera caso, o enunciado da consulta debe ser realista.

Debes ter en conta tamén que:

- Non queremos utilizar sentencias do tipo `select * from taboa where clave_primaria=valor`. Estas sentencias xa están “optimizadas” automaticamente dado que Oracle creará un índice para cada clave primaria. O mesmo ocorre se temos restricións `unique`.

- Non é necesario incluír sentencias de extrema complexidade, dado que neste traballo non se valorará o coñecemento de SQL neste sentido. Si que será común utilizar joins e/ou agrupacións e incluso algunha subconsulta, pero non será habitualmente necesario usar cláusulas máis complexas ou expresións de consulta (por suposto, non están tampouco prohibidas).

b) Indicar a solución (sentencia SELECT) a esa consulta.

c) Mostrar o plan de execución da sentencia *antes de calquera optimización da base de datos*. Debe incluírse tanto o plan de execución como a parte de predicados (*access* e *filter*, por exemplo).

Se se utiliza o programa SQL*Plus para mostrar o plan (recomendado), recordade que a indentación das operacións no plan é moi importante. Polo tanto, utilizade capturas de pantalla ou, se é texto, verificade que a indentación se mostra correctamente.

Para 2 das consultas, utiliza algún dos hints que Oracle admite, e mostra o plan de execución con ese hint. Discute as diferenzas co plan orixinal, no caso de que existan.

d) Facer os axustes da base de datos e comprobar os seus efectos.

Para cada consulta, a situación inicial debe ser a BD sen ningún tipo de optimización. A partir deste punto, fai axustes na BD para mellorar a eficiencia da sentencia da carga de traballo.

Estes axustes serán fundamentalmente crear ou eliminar índices. El algún caso quizás poderías valorar a utilización de estruturas de almacenamento diferentes das predeterminadas (IOTs, clusters etc.), pero será moito menos habitual. Debes xustificar todos os axustes que fagas.

Por suposto, é posible que un elemento como un índice concreto sirva para optimizar máis dunha consulta, pero debe validarse en cada consulta por separado.

Comproba, mostrando de novo os plans de execución, o resultado do axuste da base de datos, e discute as vantaxes e inconvenientes para cada sentencia da carga de traballo.

2. Como sabemos, a creación de índices, ou calquera outro cambio realizado no proceso de axuste, pode ter un efecto negativo nas sentencias DML.

Para verificalo, debes especificar unha sentencia DML, que poderá ser ou ben un INSERT ou ben un UPDATE, que afecte a unha das táboas que sufriu modificacións no proceso de axuste.

Ademais do plan de execución antes e despois do axuste da BD, mide o tempo de execución da sentencia (usa *set timing on* en SQL*Plus). Dado que ese tempo será moi curto para unha soa sentencia, podes lanzar un bucle PL/SQL, con algo similar a isto:

```
set timing on
begin
  for i in 1 .. 1000 loop
    Sentencia DML;
  end loop;
end;
/
```

(NOTA: Se usas un update, faino de deixos que non actualices sempre a mesma fila).

Unha vez comprobados os tempos, debes explicar razoadamente se mantés todas ou parte das optimizacións realizadas, ou eliminas algunha delas.

3. Normas de entrega e data límite

O traballo realizarase polas mesmas persoas que fixeron a primeira entrega. De cambiar a composición, debe avisarse ó profesor.

A documentación a entregar será un único documento en formato PDF, que incluírá:

- Unha portada na que conste o nome, DNI e correo electrónico das persoas que forman o grupo.
- O esquema relacional da base de datos (o feito no TT1, revisado se na corrección se detectou algún erro).
- Para cada unha das consultas de selección de datos da carga de traballo, mostra de forma agrupada o seguinte:

- O enunciado e a solución en SQL.
 - O plan de execución de Oracle para esa consulta (recorda que en 2 das consultas terás 2 plans ó utilizar hints).
 - Os axustes da base de datos, contendo as sentencias necesarias (`create index`, `alter table` etc.) e a xustificación de por que decidiu facerse cada cambio.
 - A comprobación dos axustes, vendo de novo o plan de execución e discutindo as mellorías ou empeoramento da eficiencia da consulta.
- A execución da consulta DML antes de facer ningunha optimización, e despois de facelas todas. Razóade se mantedes ou eliminades cada un dos elementos optimizados en base a esta comprobación.

A entrega realizarase usando o moodle da UDC e só por unha das persoas que forman o grupo, preferiblemente a mesma que realizou a primeira entrega.

A data límite para a entrega é o día 24/05/2024.