

AMPLIACIÓN de BASES DE DATOS

(Profesor: Héctor Gómez Gauchía)

Práctica 4 SEMANA 10 - Optimización de Consultas: Estadísticas e Índices

Resultado:

- En un archivo reúne Instrucciones, Resultados de las ejecuciones y tus respuestas a las preguntas.
- Haz lista de dudas concretas sin resolver sobre tus respuestas para consultar con el profesor

Modo de entrega:

- Si terminas alguno de los apartados indicados, avisa al profesor para puntuar la parte de participación en clase y sigue sus instrucciones.
- Los conceptos de esta práctica se evalúan en el examen Final

APARTADO 1 Preparación para esta práctica

- Ejecuta crear-BDejemplo.sql que viene en esta práctica
- Carga en Oracle las dos tablas PELISAHORA y PELISHIST como se indicó en la Práctica 4 Semana 9 (la pasada). Debes cargarlas con los mismos formatos de atributos para que las condiciones funcionen sin problemas.
- Si no lo has hecho todavía, crea sobre las dos tablas:
 - Crea las PKs sobre el atributo ID con ALTER TABLE (asígnale un nombre, como PK_pelisactual)
 - Crea un índice único sobre titulo : idx_pelisahora_unico_titulo
 - Crea un índice bitmap sobre genero idx_pelisahora__bit_genero
 - Crea un índice de función sobre round(drama) : idx_pelisahora__fun_drama

APARTADO 2 → Cómo determinar problemas de rendimiento en consultas

Para determinar **problemas de rendimiento en consultas**, además de ver el Plan de Ejecución de una consulta, es útil analizar estadísticas estimadas y reales de ejecución de consultas, así como de índices y tablas.

Sección 1.- Paquete DBMS_XPLAN: Estadísticas de ejecución real y estimaciones de una Consulta

Este paquete sirve cuando se necesitan estadísticas con más detalles que con el EXPLAIN (si quieres tener el árbol del plan, se ve más fácil con el explain y con la consulta jerárquica). Además así se obtienen datos reales de ejecución (hasta ahora eran estimados). Para ello completa los siguientes apartados:

a) Estadísticas con datos reales de ejecución

Teniendo en cuenta las columnas que se obtienen en el DBMS_XPLAN.DISPLAY_CURSOR:

- Starts : cuantas veces se ha ejecutado esa operación
- E-Rows : Filas estimadas de la operación de esa línea (antes de ejecutar)
- A-Rows : Filas reales (cuando ejecuta esa operación) de la operación de esa línea:
 - Cuando están bien estimadas las E-Rows suele coincidir con el núm. filas de E-Rows multiplicado por Starts: si hay mucha diferencia es que ese plan no va a ser eficiente.
- Buffers: memoria que ha usado para cada operación

SE PIDE:

- Ejecuta las siguientes instrucciones
- Razona las operaciones de acceso a las tablas e índices que hacen las consultas
- Contesta a las preguntas después de cada consulta

NOTA: La instrucción `/*+ GATHER_PLAN_STATISTICS */` es una "hint", para que Oracle almacene estadísticas para esa consulta mientras la ejecuta. El literal `/* prac4s10 consulta-1 */` es un comentario para darle un nombre único a la consulta y poder recuperarla luego.

```
select /*+ GATHER_PLAN_STATISTICS */          /* prac4s10 consulta-1 */
  A.ID, A.titulo, round(A.drama), H.genero
from PELISHIST H, PELISAHORA A
where A.ID = H.ID and (round(A.drama) = 43 or round(A.drama) = 50);
```

```
SET LINESIZE 130
SELECT *
FROM TABLE(DBMS_XPLAN.DISPLAY_CURSOR
              (format => 'ALLSTATS LAST'));
```

-- Esta consulta te da el plan de la última ejecución y sus estadísticas. Pero, a veces, es la propia instrucción la que toma como última. Si el plan no es de tu instrucción, prueba a repetir o mira al final de este documento donde se explica cómo solucionarlo.

RESPONDE:

Sec1.1.- Compara E-Row y A-Row. Se parecen los valores? Sí? entonces ha hecho bien la estimación?

- A veces tienes que multiplicar E-rows por Starts para tener el total
- de la operación completa ejecutada, que es lo que da las A-Rows

Sec1.2.- Observa cómo, aunque opera con pocas filas maneja muchos buffers:

- Usa muchos menos buffers en accesos a índice que a la tabla. ¿Porqué?

Sec1.3.- Esta consulta-2 tiene menos operaciones que la consulta-1:

- a.- Observando las operaciones: ¿Cual de las consultas te parece más eficiente, la anterior o esta?
- b.- Observando la A-Rows: ¿Cual de las consultas te parece más eficiente, la anterior o esta?
- c.- Observando los buffers: ¿Cual de las consultas te parece más eficiente, la anterior o esta?
- d.- Observando todo: ¿Cual de las consultas te parece más eficiente, la anterior o esta?

```
select /*+ GATHER_PLAN_STATISTICS */ /* prac4s10 consulta-2 */
  A.ID, A.titulo, A.genero
from PELISAHORA A, PELISHIST H
where A.ID = H.ID and A.ID > 100 and A.genero = 'Terror' ;
```

```
SELECT *
FROM TABLE(DBMS_XPLAN.DISPLAY_CURSOR
              (format => 'ALLSTATS LAST'));
```

Sección 2.- Características importantes de tus tablas e índices consultando USER_INDEXES y USER_TABLES

1.- Tabla user_indexes y dba_indexes : (si son índices de otro usuario trabaja con: all_indexes)

SE PIDE:

- Para entender los atributos BLEVEL, DISTINCT_KEYS, LEAF_BLOCKS, NUM_ROWS puedes consultar la explicación que vimos en clase de teoría.
- Ejecuta cada una de las instrucciones, copia el resultado de ejecutar la instrucción
- Contesta a las preguntas de cada consulta

1.a) desc user_indexes – Qué atributos parecen interesantes?

1.b) select INDEX_NAME, INDEX_TYPE, TABLE_NAME, UNIQUENESS
from user_indexes --- Índices de tu usuario
Where TABLE_NAME = 'PELISAHORA';

1.b.1 Porqué los índices de función y bitmap no son únicos?

1.c) select INDEX_NAME, INITIAL_EXTENT, BLEVEL, LEAF_BLOCKS, NUM_ROWS
from user_indexes
Where TABLE_NAME = 'PELISHIST';

1.c.1 Porqué tienen tan pocos niveles y bloques hoja estos índices?

1.c.2 Porqué el índice de género solo tiene 25 filas? demuéstalo con una consulta a la tabla

1.d) select INDEX_NAME, DISTINCT_KEYS, AVG_LEAF_BLOCKS_PER_KEY,
AVG_DATA_BLOCKS_PER_KEY, CLUSTERING_FACTOR
from user_indexes
Where TABLE_NAME = 'PELISHIST';

1.d.1 Qué son las DISTINCT_KEYS : compruébalo con una consulta a cada tabla

1.d.2 (para nota) Qué es el CLUSTERING_FACTOR?

1.e) (para nota) Para ver mayor cantidad de valores de índices con muchas más filas

- Compara LEAF_BLOCKS, NUM_ROWS de varios índices: tienen la misma relación?
¿Qué te puede indicar? (busca en la web el sentido)

```
SELECT OWNER, INDEX_NAME, BLEVEL, LEAF_BLOCKS, NUM_ROWS  
from dba_indexes -- los índices del sistema  
where NUM_ROWS > 50000;
```

2.- Tabla user_tables : (si son tablas de otro usuario trabaja con: all_tables)

SE PIDE:

- Ejecuta cada una de las instrucciones, copia el resultado de ejecutar la instrucción
- Describe qué obtienes con cada una, describe sus atributos (los más importantes en 2.c) y valores obtenidos. Necesitarás consultar documentación en la web.

2.a) desc user_tables

2.b) select table_name, num_rows, blocks, avg_row_len
from user_tables;

2.b.1 Qué son los blocks? Qué relación tienen con num_rows?

2.b.2 Porqué avg_row_len es diferente en la table pelsiahora y en pelishist?

Sección 3 : Estadísticas de una Tabla y de Columna desde el SQL DEVELOPER

SE PIDE:

- Explora estas instrucciones, ejecútalas e indica qué obtienes
 - Explica los atributos que reconoces, que creas más importantes
-
- a) Encima de una tabla botón derecho: menu contexto "Estadísticas" + “Valida Estructura”
 - b) Encima de una tabla + B.dcho: menu contexto "Estadísticas" + Recopilar Estadística (equivale al ANALIZE)
 - c) Ahora se puede consultar, un vez abierta esa tabla, en la ventana derecha con pestañas: en la pestaña "estadísticas"
 - d) Desde la misma ventana con pestañas de una tabla: pestaña "Estadística"
 - Abajo en ventana "Estadísticas de Columna"
 - e) En "Refrescar = 5" refresca datos cada 5 segundos
 - f) Desde la misma ventana con pestañas de una tabla: pestaña "Detalles"

NOTA: Por si da aviso de “clave caducada”: cambiar clave de un usuario (desde ADMINUSER)

```
alter user XXX identified CLAVENUEVA replace CLAVEANTIGUA
```

Sección 4 .- Paquete DBMS_STATS: Estadísticas del Optimizador de Consultas

(para nota)

Gestión de estadísticas de tablas e índices que usa el Optimizador para calcular costes en la decisión de escoger su Plan de Ejecución. El optimizador usa las estadísticas que tengas para decidir el plan de ejecución de las operaciones. Si no se cambian nunca, puede que estén obsoletas. Conviene obtenerlas de nuevo cada cierto tiempo.

SE PIDE:

- Ejecuta cada una de las instrucciones, copia el resultado de ejecutar la instrucción
 - Describe qué obtienes con cada una, describe sus atributos y valores obtenidos
- Necesitarás consultar la web para saber el modo de usarlo.*

a) EXEC DBMS_STATS.GATHER_INDEX_STATS('tuusuario', 'PK_PELISHIST');
(ordena que se almacenen estadísticas)

b) EXEC DBMS_STATS.GATHER_TABLE_STATS('tuusuario','PELISHIST');

c)

```
select OWNER,INDEX_NAME,NUM_ROWS,LAST_ANALYZED,
       BLEVEL,LEAF_BLOCKS,DISTINCT_KEYS
from dba_indexes
where owner = 'tuusuario'
and index_name ='PK_PELISHIST';
```

OTROS MATERIALES EXTRA:

Preparación: Si no se si tengo permisos para acceder a las tablas necesarias, puedo probar a consultarlas:

```
Select * from V$SQL_PLAN_STATISTICS_ALL ;
Select * from V$SQL;
Select * from V$SQL_PLAN;
```

Más precisión: en el apartado de obtener estadísticas /* GATHER ... */, hay una tabla donde se registra el número de consulta con el nombre que le dimos /* prac4s10 ... */ : Así se obtiene el número interno sql_id

-- necesitamos obtener valores que identifiquen la última consulta ejecutada
-- para usarlos en el siguiente paso

```
SELECT sql_id, child_number
FROM v$sql
WHERE sql_text LIKE '% miusuario secl a)1 a)1 %';
```

/* ha dado dos filas: si una no es la de tu consulta, es la de debajo:

```
sql_id , child_number
6awy4pjpmtms3 0 --> el de la consulta actual
8wt3fkb9c709x 0 */ --> el de la que quiero obtener
```

-- Luego se consulta como antes pero con el número de la instrucción:

```
SELECT * FROM TABLE(DBMS_XPLAN.DISPLAY_CURSOR
('8wt3fkb9c709x',0,format => 'ALLSTATS LAST'));
```