**Bases de Datos**

**Trabajo de Optimización de Consultas**

**Facultad de Ingeniería**

**Universidad de Antioquia**

Apoyados en la Base de Datos que cada equipo ha diseñado a lo largo del curso resuelva los siguientes interrogantes.

1. **System R.**

Utilizando el algoritmo de optimización de System-R encuentre la estrategia de menor costo para resolver una consulta de la forma:

**Select** codproducto, descripción, marca, valorventa, stockminimo

**From** productos

**Where** marca = ’Bimbo’ and

valorventa >= 15.000

TR = 14.600 , BR = 343, IR = 3.000

|  |  |
| --- | --- |
| **Esquema de Almacenamiento** | **Tabla** |
| Hash | Productos |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Índice** | **Tabla asociada** | **Tamaño** | **Tamaño en Paginas** | **Tamaño en Megas** | **Altura** | **Imagen**  **I** |
| PK\_productos | productos | 3 | 11 | 0,042 | 3 | 14.600 |
| IDX\_Productos\_marca | productos | 20 | 72 | 0,278 | 4 | 3.000 |

***Solucion:***

Opcion 1: No factible, No hay índice cluster en las condiciones de nuestra consulta.

Opcion 2: Idem 1

Opción 3: Tenemos el índice no cluster sobre marca que hace parte de la condición en el

Where. Costo : TR/IR = 14.600 / 3.000 = 5

Opcion 4: Los datos no están almacenados independientemente.

Opción 5 : Idem1

Opción 6: No aplica, no hay índice para valorventa.

Opcion 7: Tenemos el índice no cluster sobre el campo marca que hacen parte de la condición en el Where. Costo : TR = 14.600

Opción 8 : No factible, ya que se pueden cumplir las opciones 3 y 7.

*El menor costo lo da la opción 3,* donde obtenemos aproximadamente 5 tuplas que cumplen

las condiciones del Where y se asume que se encuentran en bloques diferentes. Por lo tanto la mejor estrategia seleccionada por el optimizador será la opción 3.

1. **Join entre dos tablas sin índices.**

Utilizando los métodos de clasificación-mezcla y el de creación de índices calcule el costo de un join con las siguientes características.(asumir que las tablas no tienen índice).

**Select** \*

**From** pedidos, detallepedido

|  |  |
| --- | --- |
| **Esquema de Almacenamiento** | **Tabla** |
| Hash | Pedidos |
| Hash | DetallePedido |

**Where** pedidos.nropedido = detallepedido.nropedido

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Índice** | **Tabla asociada** | **Tamaño** | **Tamaño en Paginas** | **Tamaño en Megas** | **Altura** |
| PK\_pedidos | Pedidos | 6 | 4 | 0,012 | 3 |
| PK\_nropedido\_idproducto\_nrolote | DetallePedido | 15 | 157 | 0,612 | 4 |
| IDX\_DetallePedido\_idProducto | DetallePedido | 3 | 32 | 0,122 | 3 |
| IDX\_Pedidos\_idProveedor | Pedidos | 11 | 6 | 0,022 | 3 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Método Clasificación Mezcla(merge Join)** | | | |  | | |  | |  | |  | |  | |
| **Tablas** | **Casos** | **Registro de la tabla Ts** | **Registro de la tabla TR** | | **Bloques BS** | **Bloques BR** | | **IR.S** | | **M** | | **Costo Total** | |  |
| R = pedidos, S=detallepedido | 1 | 42.800 | 2.080 | | 1.411 | 24 | | 89.024.000 | | 23 | | 9.994 | |  |
| 2 | 600.000 | 29.158,88 | | 19.775 | 6.739 | | 17.495.327.103 | | 225.151 | |  |
| 3 | 1.200.000 | 58.317,76 | | 39.551 | 13.477 | | 69.981.308.411 | | 477.250 | |  |
|  |  |  |  | |  |  | |  | |  | |  | |  |
| **Método Creación de Indices** | | | | |  |  | |  | |  | |  | |  |
| **Tablas** | **Casos** | **Registro de la tabla Ts** | **Registro de la tabla TR** | | **Bloques BS** | **Bloques BR** | | **I** | | **J** | | **M** | | **Costo Total** |
| R = pedidos, S=detallepedido | 1 | 42.800 | 2.080 | | 1.411 | 24 | | 42.800 | | 2.080 | | 23 | | 378.648 |
| 2 | 600.000 | 29.159 | | 19.775 | 6.739 | | 600.000 | | 29.159 | | 6.450.349 |
| 3 | 1.200.000 | 58.318 | | 39.551 | 13.477 | | 1.200.000 | | 58.318 | | 13.431.227 |

El costo de input para **merge Join** involucra el costo de ordenar cada una de las relaciones participantes por el atributo de junta (podría darse el caso que alguna de las relaciones ya estuviera ordenada). A esto debemos adicionarle el costo del merge, proceso en el cual se buscan en ambas relaciones ordenadas las tuplas que cumplen la condicin de junta (es decir, las tuplas que pertenecen a la misma partición), luego de lo cual se realiza la junta de estas tuplas. El costo de merge es BR + BS, dado que se recorre una vez cada relación

El método de creación por índices le suma un costo adicional al costo total, si las tablas tuvieran índices el costo total seria:





Al costo anterior debemos sumarle el costo por creación de índices: , que es un costo adicional, aumentando

los tiempos de repuesta. El crecimiento se debe al producto del tamaño de la relación por el logaritmo de de este tamaño BRlogMJ.

El crecimiento en el tamaño esperado de la reunión natural por la división de I en clasificación-mezcla y de j para creación de índices.

De los dos métodos el de mejor rendimiento es el de clasificación-mezcla ya que tiene menores E/S.

1. **Join entre dos tablas con índices.**

Utilizando join con índices evaluar la consulta del numeral dos, adicionando las siguientes

|  |  |
| --- | --- |
| **Esquema de Almacenamiento** | **Tabla** |
| Hash | Pedidos |
| Hash | DetallePedido |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Índice** | **Tabla asociada** | **Tamaño** | **Tamaño en Paginas** | **Tamaño en Megas** | **Altura** |
| PK\_pedidos | Pedidos | 6 | 4 | 0,012 | 3 |
| PK\_nropedido\_idproducto\_nrolote | DetallePedido | 15 | 157 | 0,612 | 4 |
| IDX\_DetallePedido\_idProducto | DetallePedido | 3 | 32 | 0,122 | 3 |
| IDX\_Pedidos\_idProveedor | Pedidos | 11 | 6 | 0,022 | 3 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Tablas** | **Casos** | **Registro de la tabla Ts** | **Registro de la tabla TR** | **Bloques BS** | **Bloques BR** | **I** | **J** | **Costo Total** |
| R = pedidos, S=detallepedido | 1 | 42.800 | 2.080 | 1.411 | 24 | 42.800 | 2.080 | 2.197 |
| 2 | 600.000 | 29.159 | 19.775 | 6.739 | 600.000 | 29.159 | 43.598 |
| 3 | 1.200.000 | 58.318 | 39.551 | 13.477 | 1.200.000 | 58.318 | 87.194 |

1. **Join entre tres tablas con índices.**

**Select** \*

**From** productos a, pedidos b, detallepedido c

**Where** a.codproducto = c.idproducto and

b.nropedido = c.nropedido

*Premisas :* Hay una clave foránea de tabla-1 hacia tabla-2 con los atributos

usados para construir la condición del join.

Tabla-1 tiene un índice cluster en col-i.

Tablas-2 tiene índice cluster en col-j.

Tablas-3 tiene índice no cluster en col-k.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Tablas 1 y 2** |  |  |  |  |  |  |  |  |
| **Tablas** | **Casos** | **Registro de la tabla Ts** | **Registro de la tabla TR** | **Bloques BS** | **Bloques BR** | **I** | **J** | **Costo Total** |
| R = pedidos, S=detallepedido | 1 | 42.800 | 2.080 | 1.411 | 24 | 42.800 | 2.080 | 2.197 |
| 2 | 600.000 | 29.159 | 19.775 | 6.739 | 600.000 | 29.159 | 43.598 |
| 3 | 1.200.000 | 58.318 | 39.551 | 13.477 | 1.200.000 | 58.318 | 87.194 |
|  |  |  |  |  |  |  |  |  |
| **Tabla 3 y el resultado de tabla 1 y 2** | |  |  |  |  |  |  |  |
| **Tablas** | **Casos** | **Registro de la tabla TRS** | **Registro de la tabla TU** | **Bloques BR** | **Bloques TRS** | **I** | **J** | **Costo Total** |
| RS = resultado entre R y S  U = productos | 1 | 42.800 | 14.600 | 1.891 | 645 | 42.800 | 14.600 | 89.383 |
| 2 | 600.000 | 204.673 | 26.514 | 9.044 | 600.000 | 204.673 | 1.253.027 |
| 3 | 1.200.000 | 409.346 | 53.027 | 18.089 | 1.200.000 | 409.346 | 2.506.055 |

1. **Optimización algebraica:**

Select\*

From Productos a, Pedidos b, DetallePedidos c

Where a.IdProducto = c.IdProducto and c.NroPedido = b.nroPedido

Productos

DetallePedido

Pedidos

X

X

σa.IdProducto=c.IdProducto and c.NroPedido=b.nroPedido

πIdProducto,Descripcion,Marca,NroPedido,NroFactura,UnidadesPedido,NroLote,FechaVencimiento

**Caso a**

Select Marca

From ProductosVencidos

Where FechaVencimiento < sysdate

Order by desc

πMarca

σFechaVencimiento>sysdate

ProductosVencidos

πMarca

DetallePedido

Pedidos

X

X

Productos

σFechaVencimiento>sysdate

πIdProducto,Descripcion,Marca,NroPedido,NroFactura,UnidadesPedido,NroLote,FechaVencimiento

σa.IdProducto=c.IdProducto and c.NroPedido=b.nroPedido

πDesc

πMarca

Paso 3 (diapositiva 16):

πDesc

σFechaVencimiento>sysdate

πIdProducto,Descripcion,Marca,NroPedido,NroFactura,UnidadesPedido,NroLote,FechaVencimiento

σa.IdProducto=c.IdProducto

σc.NroPedido=b.nroPedido

Productos

DetallePedido

Pedidos

X

X

Paso 4 (diapositiva 17):

πMarca

πDesc

πIdProducto,Descripcion,Marca,NroPedido,NroFactura,UnidadesPedido,NroLote,FechaVencimiento

σa.IdProducto=c.IdProducto

Productos

X

σc.NroPedido=b.nroPedido

X

DetallePedido

Pedidos

σFechaVencimiento>sysdate

Paso 5 (diapositiva 18): Paso 4 (diapositiva 17):

πMarca

πDesc

σa.IdProducto=c.IdProducto

Productos

X

σc.NroPedido=b.nroPedido

X

DetallePedido

Pedidos

σFechaVencimiento>sysdate

Paso 6 (diapositiva 19):

πMarca

πDesc

σa.IdProducto=c.IdProducto

πMarca,a.IdProducto,c.IdProducto

Productos

X

σc.NroPedido=b.nroPedido

X

DetallePedido

Pedidos

σFechaVencimiento>sysdate

πMarca

Paso 7 (diapositiva 20):

πDesc

σa.IdProducto=c.IdProducto

πMarca,a.IdProducto

X

πc.IdProducto

σc.NroPedido=b.nroPedido

X

DetallePedido

Pedidos

σFechaVencimiento>sysdate

Productos

Paso 8 (diapositiva 21):

πMarca

πc.NroPedido,c.IdProducto

π b.nroPedido

X

σc.NroPedido=b.nroPedido

πc.IdProducto

X

πMarca,a.IdProducto

σa.IdProducto=c.IdProducto

Productos

DetallePedido

σFechaVencimiento>sysdate

Pedidos

πDesc

**Caso b**

Select Marca

From ProductosVencidos

Where FechaVencimiento < sysdate

πMarca

σFechaVencimiento>sysdate

ProductosVencidos

DetallePedido

Pedidos

X

X

Productos

πMarca

σFechaVencimiento>sysdate

πIdProducto,Descripcion,Marca,NroPedido,NroFactura,UnidadesPedido,NroLote,FechaVencimiento

σa.IdProducto=c.IdProducto and c.NroPedido=b.nroPedido

Paso 3 (diapositiva 16):

πMarca

σFechaVencimiento>sysdate

πIdProducto,Descripcion,Marca,NroPedido,NroFactura,UnidadesPedido,NroLote,FechaVencimiento

σa.IdProducto=c.IdProducto

σc.NroPedido=b.nroPedido

Productos

DetallePedido

Pedidos

X

X

Paso 4 (diapositiva 17):

πMarca

πIdProducto,Descripcion,Marca,NroPedido,NroFactura,UnidadesPedido,NroLote,FechaVencimiento

σa.IdProducto=c.IdProducto

Productos

X

σc.NroPedido=b.nroPedido

X

DetallePedido

Pedidos

σFechaVencimiento>sysdate

Paso 5 (diapositiva 18):

πMarca

σa.IdProducto=c.IdProducto

Productos

X

σc.NroPedido=b.nroPedido

X

DetallePedido

Pedidos

σFechaVencimiento>sysdate

Paso 6 (diapositiva 19):

πMarca

σa.IdProducto=c.IdProducto

πMarca,a.IdProducto,c.IdProducto

Productos

X

σc.NroPedido=b.nroPedido

X

DetallePedido

Pedidos

σFechaVencimiento>sysdate

Paso 7 (diapositiva 20):

πMarca

σa.IdProducto=c.IdProducto

πMarca,a.IdProducto

X

πc.IdProducto

σc.NroPedido=b.nroPedido

X

DetallePedido

Pedidos

σFechaVencimiento>sysdate

Productos

Paso 8 (diapositiva 21):

πc.NroPedido,c.IdProducto

π b.nroPedido

X

σc.NroPedido=b.nroPedido

πc.IdProducto

X

πMarca,a.IdProducto

σa.IdProducto=c.IdProducto

Productos

DetallePedido

σFechaVencimiento>sysdate

Pedidos

πMarca