

Base de Datos (75.15 / 75.28 / 95.05)

Evaluación Integradora - 4 de julio de 2018

TEMA 20181C1						Padrón: _____
ACR		Proc.		Esp.		Apellido: _____
Seg.		CyT		DW		Nombre: _____
Corrigió:						Cantidad de hojas: _____
Nota:						<input type="checkbox"/> Aprobado <input type="checkbox"/> Insuficiente

Criterio de aprobación: El examen está compuesto por 6 ítems, cada uno de los cuales se corrige como B/B-/Reg/Reg-/M. Se aprueba con nota mayor o igual a 4(cuatro), equivalente a desarrollar el 60 % del examen correctamente.

1. (*Álgebra y Cálculo Relacionales*) Una empresa aseguradora que ofrece polizas para distintos tipos de vehículos (automóviles, barcos y aeronaves) mantiene las siguientes relaciones que guardan información sobre sus clientes y los vehículos que poseen cubiertos:

- `Clientes(CUIT, nombre_cliente, domicilio, teléfono)`
- `Vehículos(id_vehículo, CUIT_cliente, fecha_alta)`
- `Automóviles(id_vehículo, nro_chasis, patente, marca, modelo)`
- `Barcos(id_vehículo, nro_casco, astillero)`
- `Aeronaves(id_vehículo, matrícula, marca, modelo)`

Se pide:

- a) Escriba una consulta en Álgebra Relacional que muestre el CUIT, nombre y teléfono de los clientes que poseen tanto algún barco como alguna aeronave.
- b) Escriba una consulta en Cálculo Relacional de Tuplas (C.R.T.) que muestre el CUIT, nombre y teléfono de los clientes que no poseen ningún barco.

Nota: La relación entre *Vehículos* y (*Automóviles*, *Barcos*, *Aeronaves*) es una unión.

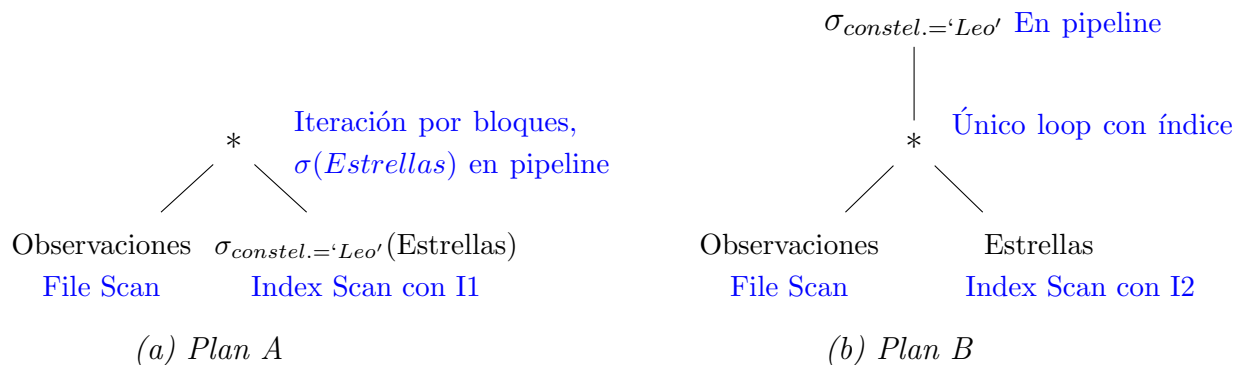
2. (Procesamiento de Consultas) La UFA (Unión de Fanáticos de la Astronomía) dispone de una base de datos con información sobre 10000 estrellas y 50000 observaciones realizadas por sus socios. La base de datos cuenta con las siguientes tablas:

- Estrellas(nombre_estrella, constelación, declinación, ascensión)
- Observaciones(nro_socio, fecha, nombre_estrella, velocidad, intensidad)

También se cuenta con los siguientes dos índices de tipo árbol:

- I1(constelación, Estrellas): Índice de *clustering* (agrupamiento) por el atributo constelación en la tabla Estrellas.
- I2(nombre_estrella, Estrellas): Índice secundario por el atributo clave nombre_estrella en la tabla Estrellas.

Con el objetivo de encontrar todas las observaciones de estrellas que pertenecen a la constelación de 'Leo', el SGBD construye los siguientes dos planes de ejecución:



Como puede observar, el *Plan A* respeta las reglas de optimización algebraica para realizar la selección –utilizando el índice correspondiente– previo a la junta –que se realiza sin índices–, mientras que el *Plan B* realiza primero la junta con índice, y luego realiza la selección.

Se pide:

- a) Estime el costo de cada plan en términos de cantidad de accesos a disco, y determine cuál de los dos es más conveniente en este sentido.
- b) Proponga la construcción de un índice adicional que permita planificar esta consulta de manera de utilizar dos índices, y grafique el plan de ejecución correspondiente.

Puede utilizar para sus cálculos la siguiente información de catálogo:

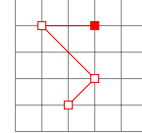
ESTRELLAS	OBSERVACIONES
$n(\text{Estrellas}) = 10000$	$n(\text{Observaciones}) = 50000$
$B(\text{Estrellas}) = 1000$	$B(\text{Observaciones}) = 5000$
$V(\text{constelación}, \text{Estrellas}) = 10$	$V(\text{nombre_estrella}, \text{Observaciones}) = 10000$
$H(I1(\text{constelación}, \text{Estrellas})) = 1$	
$H(I2(\text{nombre_estrella}, \text{Estrellas})) = 4$	

Nota 1: Considere la utilización de sólo un bloque de memoria por cada tabla/índice.

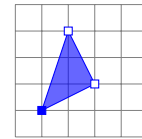
Nota 2: No considere el costo de almacenar el resultado final de la consulta en disco.

3. (*Bases de datos espaciales*) Disponemos de dos tablas con la traza de todas las rutas nacionales y los límites provinciales de la República Argentina. Dichas tablas fueron creadas en *Postgres/PostGIS* de la siguiente forma:

```
CREATE TABLE RutasNacionales (
  nro_ruta INT PRIMARY KEY,
  traza GEOGRAPHY(LINESTRING)
);
```



```
CREATE TABLE Provincias (
  nombre_provincia VARCHAR(30) PRIMARY KEY,
  territorio GEOGRAPHY(MULTIPOLYGON)
);
```



El atributo *traza* de la tabla *RutasNacionales* almacena objetos geométricos de tipo *LineString* (cada objeto es un conjunto de segmentos consecutivos), mientras que el atributo *territorio* de la tabla *Provincias* almacena objetos de tipo *MultiPolygon* (cada objeto es un conjunto de polígonos).

Escriba una consulta SQL que, a partir de estas dos tablas, encuentre el número de la ruta que atraviesa la mayor cantidad de provincias en el país, indicando el número de la ruta, su longitud total y la cantidad de provincias que atraviesa.

Como referencia puede considerar la siguiente consulta SQL que encuentra, de entre todas las provincias atravesadas por la Ruta 3, aquella de mayor superficie, y muestra el nombre de la provincia y la cantidad de kilómetros de la Ruta 3 que la atraviesan.

```
SELECT nombre_provincia,
       ST_Length(ST_Intersection(p.territorio, r.traza))/1000 AS km
FROM Provincias p, Rutas r
WHERE ST_Intersects(p.territorio, r.traza)=TRUE
AND r.nro_ruta=3
AND ST_Area(p.territorio) =
  (SELECT MAX(ST_Area(p2.territorio))
   FROM Provincias p2, Rutas r2
   WHERE r2.nro_ruta=3
   AND ST_Intersects(p2.territorio,r2.traza)=TRUE);
```

Nota: Puede suponer, si lo desea, que existe una única ruta que atraviesa la mayor cantidad de provincias.

4. (*Seguridad*) Explique en qué consisten los ataques de *SQL injection*, y mencione dos precauciones que podría tomar para evitarlos.

5. (*Concurrencia y transacciones*) Se dice que una transacción T_1 realiza una Lectura no Repetible ó *Unrepeatable Read* cuando la misma lee un ítem X , luego otra transacción T_2 escribe ese mismo ítem, y posteriormente T_1 vuelve a leer el ítem, encontrando un valor distinto al anteriormente leído. Indique si las siguientes afirmaciones sobre la Lectura no Repetible son verdaderas ó falsas. Justifique cada una de sus respuestas.
- a) En un solapamiento recuperable puede ocurrir una Lectura no Repetible.
 - b) En un solapamiento que evita *rollbacks* en cascada puede ocurrir una Lectura no Repetible.
 - c) Aplicando el Protocolo de Lock de Dos Fases (2PL) puede ocurrir una Lectura no Repetible.
 - d) Bajo el nivel de aislamiento *Read Committed* definido en el estándar SQL puede ocurrir una Lectura no Repetible.
6. (*Data Warehousing*) Explique cuál es el papel de los procesos ETL (*Extract-Transform-Load*) en la construcción de un Data Warehouse.