

# Recuperatorio Bases de Datos: Consultas y Ecosistema

24 de Noviembre 2022

Nombre: Ariel Agustín Alvaro

Ejercicio 1: Resolver:

a) Sea la siguiente base de datos de fotografías:

usuarios (uid, nombre)

fotos (fid, dueño, tamaño)

comentarios (cid, autor, foto, texto)

Observar que *dueño* y *autor* son usuarios (*uid*), *foto* es una *fid*.

Considerar la consulta:  $\prod_{fid, uid} (fotos \bowtie \sigma_{nombre='Luis Perez'} (usuario))$ .

Asumir que hay índice primario en *nombre* para usuario, que usuario tiene 100.000 tuplas, que un bloque tiene 4096 bytes, que hay 10.000 nombres diferentes de usuarios, que un mismo nombre no se repite más de 15 veces en usuario.

1. Calcular la cantidad de claves de búsqueda por nodo del índice (árbol B+)
2. Calcular la altura del árbol B+.
3. Indicar el operador físico que conviene usar para la selección y estimar el costo en número de transferencias de bloques para este operador físico.

b) Sea la BD de una pizzería con las siguientes tablas:

cliente (cid, nombre, teléfono, dirección, edad)

pizzas (zid, nombre, tamaño, precio)

pedido (pid, cid, phora, paño, pmes, pdía, cantidad)

Considerar: cliente  $\bowtie$  pizzas  $\bowtie$  pedido

Se tiene la siguiente información de las tablas:

cliente	pizzas	pedido
500 tuplas	40 tuplas	5000 tuplas
	$V(pizzas, tamaño) = 4$	$V(pedido, zid) = 40$
	$V(pizzas, znombre) = 10$	$V(pedido, cid) = 500$
		$V(pedido, paño) = 10$

$$V(pizzas, zid) = 40$$

$$V(pedido, zid) = 5000$$

$$cliente - pizza = 100, 40, 71$$

Se considera optimización de costo usando programación dinámica.

Se cuenta con la siguiente tabla para el caso  $n = 2$ :

	{cliente, pizzas}	{cliente, pedido}	{pizzas, pedido}
Tamaño	20000	5000	5000
Costo	0	0	0
Mejor plan	Pizzas $\bowtie$ cliente	Cliente $\bowtie$ pedido	Pizzas $\bowtie$ pedido

Para {cliente, pizzas, pedido} encontrar mejor plan, su tamaño y su costo.

Ejercicio 2: Responder:

- Sea el esquema universal EmpresaTransporte = (nombre chofer, legajo chofer, nombre azafata, legajo azafata, fecha de viaje, patente de colectivo, tipo de colectivo, origen, destino, paradas intermedias). Se pide:
  - Dar ejemplo de redundancia de información y dependencia funcional que la captura.
  - Descomponer el esquema EmpresaTransporte eliminando esa redundancia de información.
- Considerar el esquema  $R = (A, B, C, D, E)$  con el conjunto de dependencias funcionales  $F = \{B \rightarrow EA, C \rightarrow E, DA \rightarrow B, E \rightarrow AC\}$ . ¿Está (B, E, A) en FNBC? Justifique su respuesta.

Ejercicio 3: Explique cómo los lenguajes de bases de datos orientadas a objetos y los lenguajes orientados a objetos difieren en cuanto a persistencia, y ponga un ejemplo en que se muestre el funcionamiento de una base de datos orientada a objetos ilustrando su relación con la persistencia

Ejercicio 4: Explique por qué OLAP es más rápido respondiendo a queries que OLTP, glosando el siguiente texto:

*"OLAP is fundamentally for read-only data stores. Classic OLAP is a Data Warehouse or Data Mart and we work with either as an OLAP cube. Conceptually you can think of an OLAP Cube like a huge Excel PivotTable. That is a structure with sides (dimensions) and data intersections (facts) that has NO JOINS.*

*The data structure is one of the reasons that OLAP is so much faster to query than OLTP. Another reason is the concept of aggregations, which are stored intersections at a level higher the leaf (bottom). An example would be as follows:*

*You may load a cube with the facts about sales (i.e. how much in dollars, how many in units, etc..) with one row (or fact) for each sales amount by the following dimensions - time, products, customers, etc..The level at which you load each dimension, for example sales by EACH day and by EACH customer, etc...is the leaf data. Of course you will often want to query aggregated values, that is sales by MONTH, by customers in a certain CITY, etc...*

*Those aggregations can be calculated at query time, or they can be pre-aggregated and stored at cube load. At query time, OLAP cubes use a combination of stored and calculated aggregations. Unlike OLTP indexes, PARTIAL aggregations can be used.*

*In addition to this, most OLAP cubes have extensive caching set up by default and most also allow for very granular cache tuning (pre-loading)."*

Ejercicio 5: Explique dos ventajas y dos desventajas de una base de datos basada en blockchain.