

BD2

TAA

1er Parcial

Tema 1 – 30 p. – Explique:

- a- Las medidas de rendimiento de los discos magnéticos. (6p)
- b- Los enfoques de organización de archivos un SGBD. (6p)
- c- Los tres tipos de índices estudiados y cuando conviene aplicarlos. (6p)
- d- Los pasos para el Procesamiento de Consultas. Haga el diagrama. (6p.)
- e- El concepto de transacción y las propiedades de las mismas que debe ser aseguradas por el SGBD (6p.)

← No entra creo

a) Medidas de rendimiento de los discos

- tiempo de acceso: tiempo desde que se hizo la solicitud hasta que se empiezan a transferir datos
 - tiempo en que se tarda en leer/escibir un dato al disco
 - * es la suma del t.de busq. y lat. rotacional (lo que tarda el brazo en llegar y el disco en girar)
- tiempo de búsqueda: tiempo que tarda el brazo en llegar a la pista del dato (esperar que el brazo se ubique)
- latencia rotacional: tiempo que tarda el disco en girar hasta que el dato esté bajo el brazo
- Velocidad de transferencia de datos: velocidad de lectura/escritura, datos por unidad de tiempo
- Tiempo medio entre fallos: Cantidad esperada de tiempo que el disco funcione sin fallos
 - tiempo hasta que halla fallos?

b) Enfoques de organización de archivos en un SGBD

? Registros de longitud fija: cada registro tiene la misma cantidad de bytes

• Registros de longitud variable: puede guardar múltiples tipos de registros, o registros con campos variables

• Secuencial: almacenados en forma ordenada de acuerdo a una clave

heap: Se guardan en cualquier lugar donde haya espacio, rápido de guardar

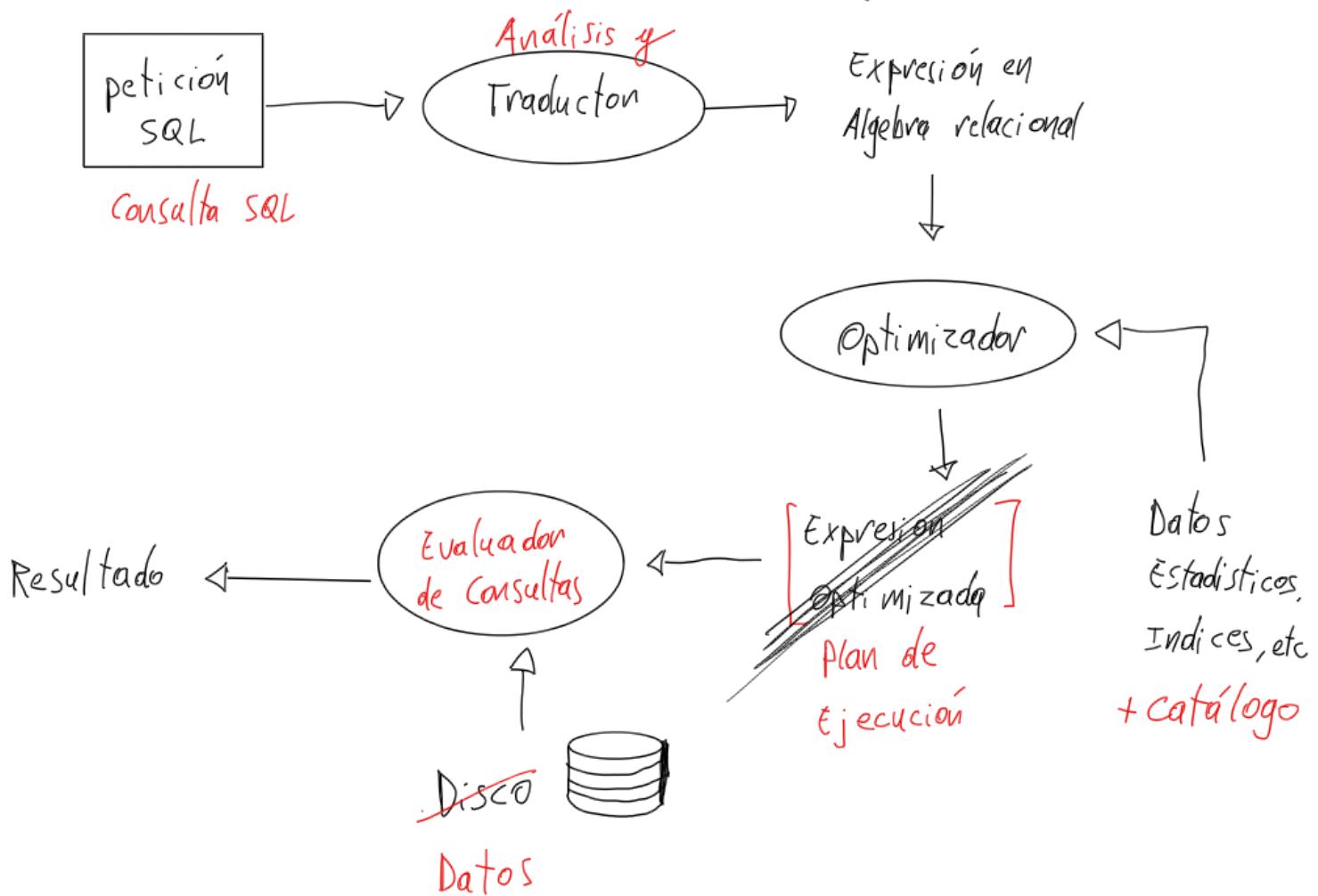
hash: Los registros se guardan en base a una función hash aplicada sobre la clave

Organización en Agrupación: Permite guardar varias tablas en un mismo archivo

c) Los tres tipos de índices estudiados y cuando conviene aplicarlos
según su implementación

- Indices Ordenados: los valores están ordenados de forma secuencial en relación a los valores de la tabla, conviene cuando realizamos búsquedas con comparación
- Indices Asociativos: los valores están distribuidos a base de una función hash, es una (Hash) buena estrategia cuando hacemos búsqueda por igualdad
- Indices Bitmap: Requiere que las filas de la tabla tengan valores discretos, es un array de bits en que sus filas corresponden a los valores discretos, y sus columnas corresponden a las filas de la tabla, es muy bueno para ahorrar espacio

d-) Los pasos para el procesamiento de consultas. Haga el diagrama



Análisis y Traducción: Hace lo mismo que un analizador léxico del compilador revisa la sintaxis y verifica con el catálogo que todas las variables introducidas correspondan a objetos que existen en la BD

Optimización:

Tema 2 - 15 p. - La siguiente tabla corresponde al archivo de la relación cliente, en el que cada bloque del archivo

corresponde a 1 fila, se pide:

- Construir un índice en forma de árbol B+ con nodos de 4 punteros para la clave primaria id, suponiendo que los registros fueron insertados según el orden de los valores de la columna nombre.
- Construir un índice hash estático cerrado con cajones de 3 elementos cuya función de asociación es " $x \bmod 5$ " sobre la columna saldo, siendo x el valor de cada fila en dicha columna.
- Explique a dónde apuntan los registros índices para caso anterior.

id	nombre	10	saldo
1	Preston Schwartz	5	282
2	Cathleen Steele	11	159
3	Tatyana Russo	7	367
4	Libby Madden	8	431
5	Orla Reid	18	317
6	Vivian Cherry	5	361
7	Kirk Jensen	1	317
8	Amanda Macias	9	190
9	Barry Morris	6	338
10	Lee Lopez	4	437
11	Elliott Fowler	9	56
12	Paula Johns		190

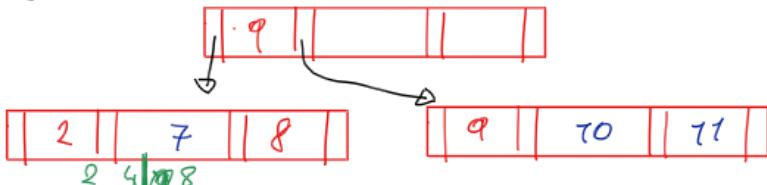
orden = 8
 --- 8
 --- 9
 --- 2
 --- 11
 --- +
 --- 10
 --- 4
 --- 5
 --- 12
 --- 1
 --- 3
 --- 6

a) $f_r = 4$

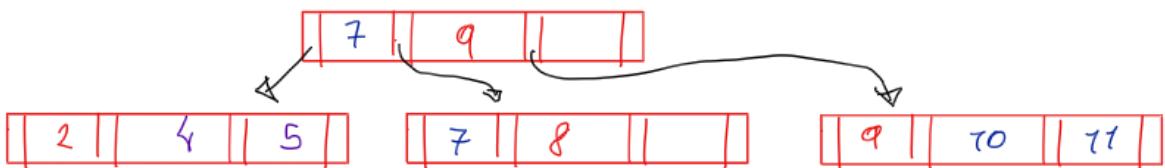
1º insert 8, 9, 2



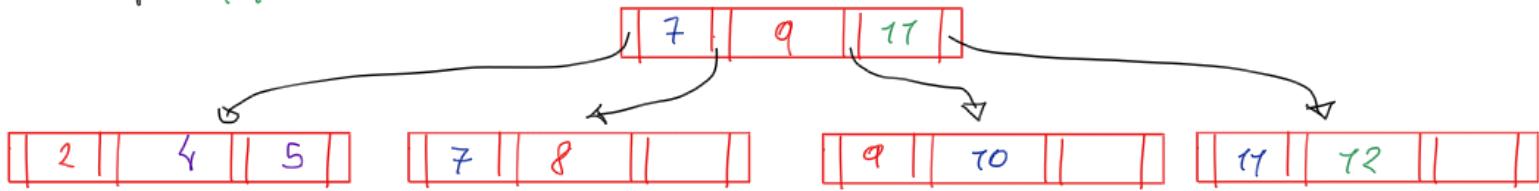
2º insert 7, 7, 10



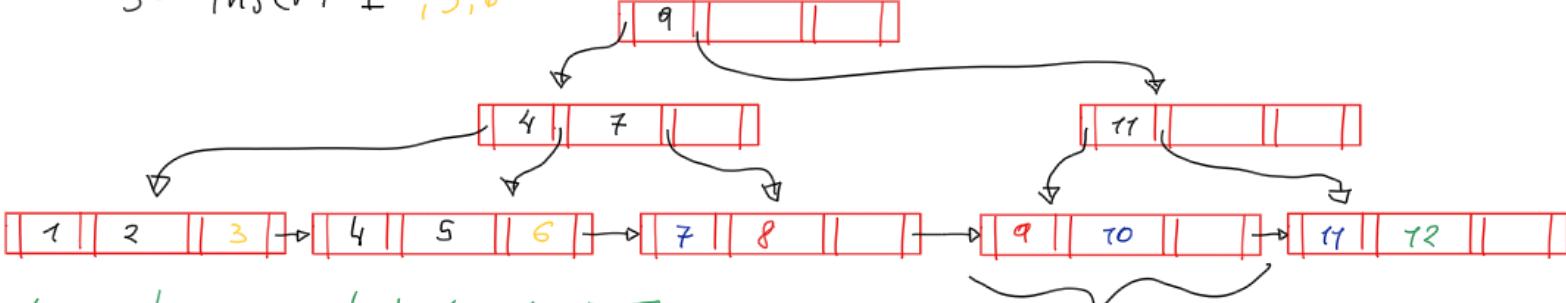
3º insert 4, 5



4º insert 12



5º insert 1, 3, 6



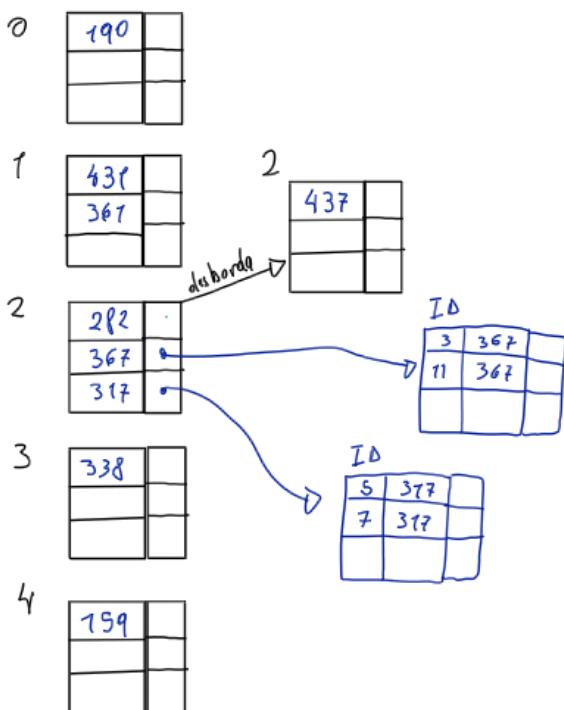
En insert 1 - se duplicó raión [235] a [1245]

- se duplicó padre [7917] a [47] ↓ [11]

9 subió a padre

b- registros fueron insertados según el orden de los valores de la columna nombre.
Construir un índice hash estático cerrado con cajones de 3 elementos cuya función de asociación es " $x \bmod 5$ " sobre la columna saldo, siendo x el valor de cada fila en dicha columna.

fr: 3



id	nombre	saldo
1	Preston Schwartz	282
2	Cathleen Steele	159
3	Tatyana Russo	367
4	Libby Madden	431
5	Orla Reid	317
6	Vivian Cherry	361
7	Kirk Jensen	317
8	Amanda Macias	190
9	Barry Morris	338
10	Lee Lopez	437
11	Elliott Fowler	367
12	Paula Johns	190

$\bmod 5$

c.) a los registros de la tabla o al cajón de punteros

Tema 3 - 10 p. - Dadas las relaciones $r1(A, B, C)$ y $r2(C, D, E)$ con las siguientes propiedades: $r1$ tiene 40.000 tuplas, $r2$ tiene 90.000 tuplas, 100 tuplas de $r1$ caben en un bloque y 45 tuplas de $r2$ que caben en un bloque. Estímese el número de accesos a bloques requeridos utilizando las siguientes estrategias para la reunión " $r1 \text{ join } r2 \text{ on } (r1.C = r2.C)$ ":

- a- Reunión en bucle anidado.
- b- Reunión en bucle anidado por bloques.
- c- Reunión por mezcla, suponiendo que la relación $r1$ debe ser ordenada externamente disponiéndose de 3 bloques de memoria intermedia.
- d- Reunión en bucle anidado indexada, suponiendo que existe un índice privado de bloques en la columna C de la relación $r2$ con nodos de 60 punteros.

si caben solo 1 de c/u en Mem.

$$n_1 = 40.000 \quad n_2 = 90.000$$

$$b_1 = \frac{40000}{100} = 400 \quad b_2 = \frac{90000}{45} = 2000$$

a.) $n_1 \times b_2 + b_1 = 40000 \times 2000 + 400 = 80.000.400 \text{ Acc. a. Bl.}$

si caben todos en Mem

$$b_1 + b_2 = 400 + 2000 = 2400 \text{ Acc. a. Bl.}$$

$$400 \cdot 2000 + 400 \\ b_1 \times b_2 + b_1 = 800.400$$

b.) $b_1 + b_2 = b_1 + b_2$

si caben en Mem = $b_1 + b_2$

c.) $M = 3$

$$b_1 = 400$$

8

$$\text{coste ordenación} = b_1 \times (2 * \lceil \log_{M-1}(b_1/M) \rceil + 1) = 6800$$

$$\text{coste Mezcla} = b_1 + b_2 = 2400$$

$$\text{Total} = 9200$$

$$\lceil \log_{60/2} (V(A,r)) \rceil \\ AA_i + 1$$

d.) $g_i = 60$

$$\text{costo} = b_r + n_r \times \text{costo índice}$$

$$400 + 40.000 \times (\lceil \log_{60/2} (90.000) \rceil + 1) = 200.400$$

3.35
asumo que C
es la columna índice de r2?

Tema 4 - 5 p. - Dada una operación de selección cualquiera basada en una condición de igualdad, explique en qué casos un SGBD utilizará los siguientes algoritmos e indique cuál es el costo asintótico de los mismos.

- 1- Búsqueda Lineal.
- 2- Búsqueda Binaria.
- 3- Búsqueda en índice primario para un atributo clave.
- 4- Búsqueda en índice secundario para atributo no clave.

1.) Busqueda Lineal: Puede utilizarse en cualquier caso pero no es preferible
el costo es la cantidad de bloques = b_r

2.) Busqueda Binaria: Puede utilizarse en el caso que la clave de búsqueda sea
la que ordena la tabla , costo = $\lceil \log_2(b_r) \rceil + \underbrace{CS(A,r)}_{\text{en caso atr. no candidato}} / f_r - 1$

3.) Busq. Ind. Primario : Cuando tenemos un índice sobre una clave que ordena la tabla
y que no tiene atributos repetidos , costo = $AA_i + 1$

4.) Busq. Ind. Sec. Atr. no clave: Cuando tenemos un índice sobre un atributo que
no está ordenado y puede contener elementos repetidos

$$\text{costo} = AA_i + 1 + CS(A,r)$$

Tema 5- 10 p. - Dadas las relaciones A(a1, a2, ..., a20), B(b1, b2, ..., b12) y C(c1, c2, ..., c15), y
la siguiente consulta:

```
select distinct A.a1, C.c1 from A join B (A.a2 = B.b3)
join C (C.c2 = B.b4) where A.a1 > 10 and B.b1 = 50;
```

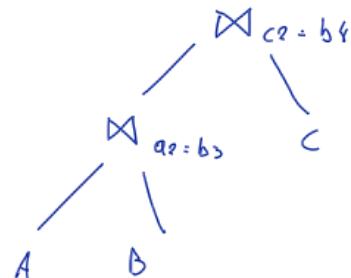
Muestre:

1. Su traducción directa al álgebra relacional.
2. Los pasos para llegar a la expresión equivalente que minimice la cantidad de datos procesado por cada operación.
3. El árbol de evaluación de la expresión final.

$$1.) \Pi_{A.a1, C.c1} \left(\sigma_{A.a1 > 10 \wedge B.b1 = 50} ((A \bowtie B) \bowtie C) \right) \quad 3.)$$

$$\Pi_{A.a1, C.c1}$$

$$\sigma_{A.a1 > 10 \wedge B.b1 = 50}$$



(Ni idea si está bien)

Tema 6 - 5p - Determine si las siguientes planificaciones son o no secuenciables en cuanto a conflictos e indique cuál es el orden de las transacciones en caso afirmativo.

P1	T1	T2
Leer(A)	Escribir(A)	
Escribir(A)	Escribir(A)	

P2	T1	T2
Leer(A)		
	Escribir(B)	Escribir(A)

P3	T1	T2	T3
Escribir(A)			
	Escribir(B)		
		Leer(C)	Escribir(C)

Si

No entra creo

Tema 7 - 5p - Determine si las siguientes planificaciones son o no secuenciables en cuanto a vistas e indique cuál es el orden de las transacciones en caso afirmativo.

P1	T1	T2
Leer(A)	Escribir(A)	
Escribir(A)	Escribir(A)	

P2	T1	T2
Leer(A)	Escribir(B)	
Leer(B)	Escribir(A)	

P3	T1	T2	T3
Leer(A)	Leer(B)	Escribir(B)	
Escribir(A)		Leer(A)	Escribir(A)

Si

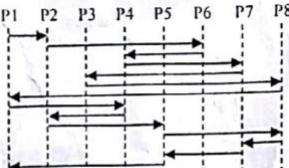
100

0.5Cp

Tema 8 - 10 pestañas

Figura que muestra la planificación de lectura del disco en un SGBD que no utiliza el algoritmo del ascensor. Se pide:

- Explique el algoritmo del ascensor y el porqué de su utilidad.
- Muestre una planificación basada en el uso del algoritmo del ascensor suponiendo que:
 - El recorrido del peine empieza en la pista 1,
 - Se pueden atender varias solicitudes por parada.
 - Las solicitudes llegan cada 1 ms.
 - Las lecturas demoran 0,5 ms por solicitud.
 - El peine se mueve a una razón de 2 pistas/ms
- Hallar la tasa de atención de solicitudes por unidad de tiempo sin y con la aplicación del algoritmo del ascensor.

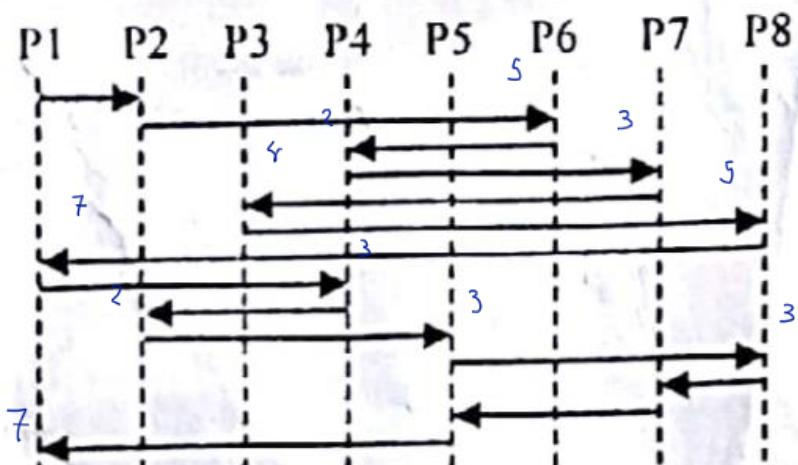


a-) **Algoritmo del Ascensor:** Es un algoritmo que se encarga de planificar el movimiento del peine para maximizar la cantidad de solicitudes atendidas por unidad de tiempo. Ayuda a maximizar el tiempo de respuesta

Planifica el movimiento del peine para que, si ya está yendo en una dirección, continúe hasta leer todas las solicitudes que están hacia esa dirección.

Si aparecen solicitudes para las cuales el peine tiene que retroceder, se reordenan esas solicitudes para atenderlas a la vuelta.

Hace que el movimiento del peine sea controlado, constante



$$\text{recorrió: } 48 \text{ pistas} = 22 \text{ ms}$$

$$\text{leyó: } 14 \text{ solicitudes} = 7 \text{ ms}$$

$$\underline{\text{total}} = 29 \text{ ms}$$

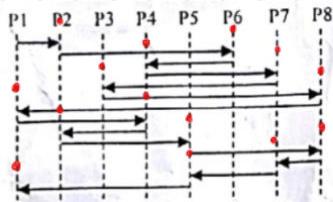
$$\text{tasa} = \frac{1}{29} = \boxed{0,034 \text{ sol./ms}}$$

Sin ascensor

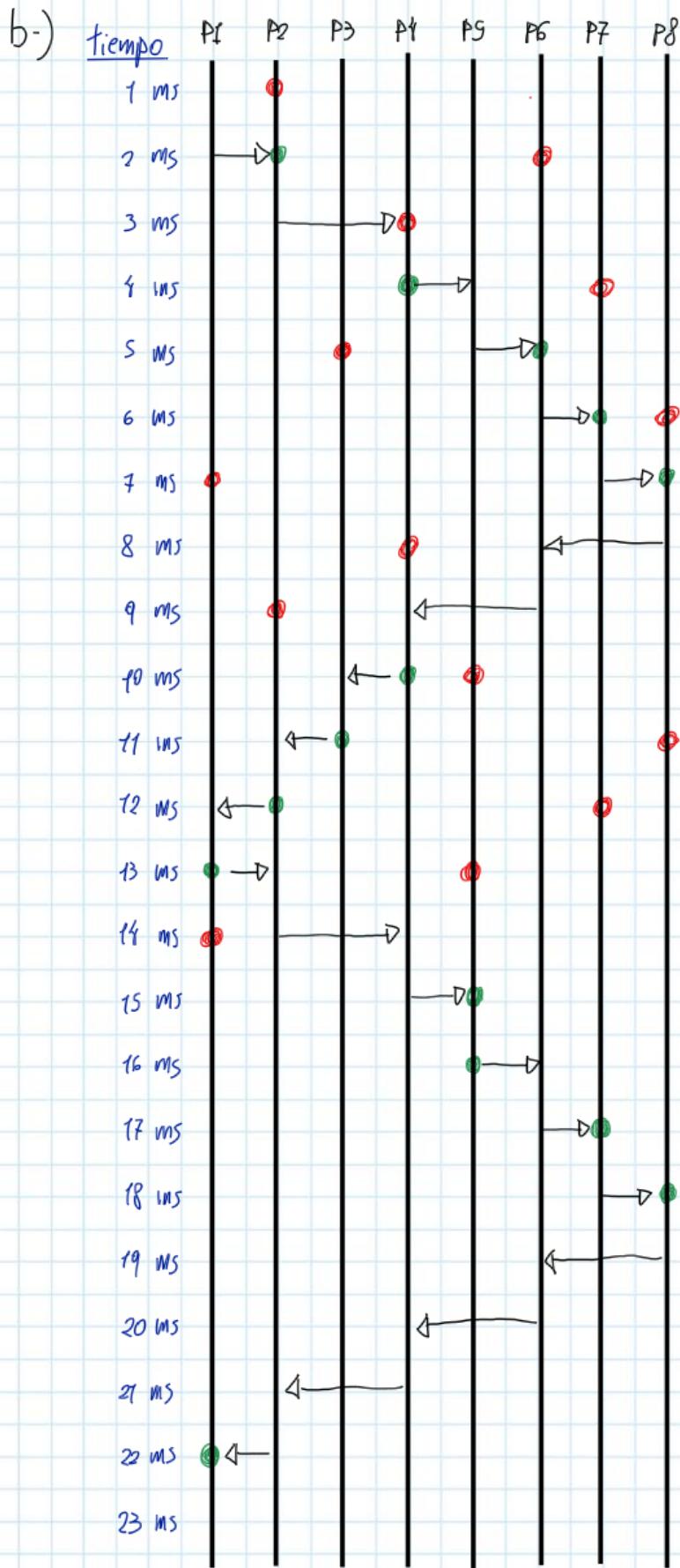
orden Solicituds

óptimo del ascensor. Se pide:

- Explique el algoritmo del ascensor y el porqué de su utilidad.
- Muestre una planificación basada en el uso del algoritmo del ascensor suponiendo que:
 - El recorrido del peine empieza en la pista 1,
 - Se pueden atender varias solicitudes por parada.
 - Las solicitudes llegan cada 1 ms.
 - Las lecturas demoran 0,5 ms por solicitud.
 - El peine se mueve a una razón de 2 pistas/ms.



2
6
4
7
3
8
1
4
2
5
8
7
5
1



$$\text{recorri\'o} = 7 + 7 + 7 + 7 = 28 \text{ pistas}$$

$$= 14 \text{ ms}$$

$$\text{leyo'} \frac{14}{28} \text{ sd} = 7 \text{ ms}$$

$$\underline{\text{Total}} = 21 \text{ ms} \quad \underline{\text{con ascensor}}$$

$$\text{Tasa} = \frac{14}{21} = 0,67 \text{ Sol./ms}$$

- ✓ Explique resumidamente y en sus propios términos a que corresponde un índice, a que un archivo índice y a que un registro índice.

Índice:

Archivo Índice:

registro índice

- ✓ Explique en sus términos a que corresponde un índice y a que un archivo índice. Así mismo enumere al menos tres tipos de índices estudiados. 5/5

Ventajas = un raid te asegura contra fallas

Raid te da una tasa de transferencia mejorada

RAID 0 : Los discos están distribuidos entre sí pero sin redundancia

RAID 1 : Por cada disco creas una copia

Raid 2 :

Raid 3 :

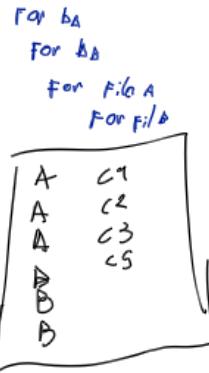
Raid 4 :

RAID 5 : Los bits de paridad están distribuidos en los N bloques

RAID 6 : Es como el Raid 5 pero guarda información adicional para protección contra fallos

Además, deben de poder plantear ejercicios hipotéticos, tales como:

1. Dadas dos tablas A y B, de Na bloques y Nb bloques, y un buffer de memoria de M bloques, muestre el estado final de la memoria luego de la 3ra iteración del ciclo mas interno del algoritmo de bucle anidado por bloques para: (a) la estrategia MRU, (b) la estrategia LRU.



Unidad de Almacenamiento

S
Sucursal

B1	B2		

C
Cliente

B1	B2	B3		
B4	B5			
				B6

LRU



M
Memoria

este
está
clavado

B1	B2	B3	B4
	B2	B3	B4

↑
MRU

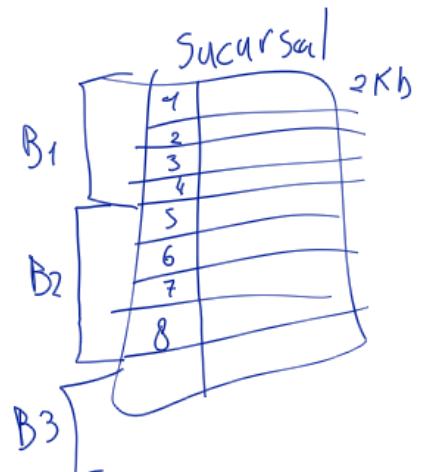
los que se desclavan son los
que se terminaron de leer
en este caso por cada fila hago
una lectura completa de C

más recientemente utilizado

```

for fila en suc
  for fila en clien
    si cli.suc = filas.now
  }
}
  
```

2. Dada la siguiente planificación de lectura de sectores del disco, muestre como la misma quedaría reordenada si se aplica el algoritmo del ascensor.



3. Dadas dos tablas A y B, de N_A bloques y N_B bloques, un buffer de memoria de M bloques, estímese el costo de la siguiente consulta SQL de selección, (a) según el algoritmo de búsqueda lineal, (b) según el algoritmo de búsqueda binaria, (c) según el algoritmo selección con índice primario, (d) según el algoritmo selección con índice secundario, ...

21:58 ✓

Tema 8 (10 p.). Dada la consulta `select * from A join B on A.a = B.b`, teniendo en cuenta que:

- La tabla A se encuentra almacenada en 20 bloques,
- La tabla B se encuentra almacenada en 15 bloques,
- La memoria cuenta con actualmente con 10 bloques libres para evaluar la consulta, y
- Que la consulta está planificada para ser evaluada conforme el algoritmo de Bucle Anidado por Bloques,

Calcule cuál será el coste de evaluación de la consulta si:

1. La estrategia de reemplazo de bloques en la memoria es LRU.
2. La estrategia de reemplazo de bloques en la memoria es MRU.

for bloq A
For bloq B
For fila A
For fila B

A

A ₁	A ₂	B ₃	B ₄	
				B ₂₀

B

B ₁	B ₂	B ₃	B ₄	B ₅
				B ₁₅

$$b_{rA} = 2^0$$

$$b_{rB} = 1^5$$

$$M = 1^0$$

↓ LRU

$$b_{AK} \left(b_B - (M - 1) \right) + b_A$$

M

A ₂	B ₁	B ₃	B ₄	
				B ₂₀

LRU

↑ MRU

MRU

M - 1

B _{A1}	1	2	3	13
12	15	16	17	18

A	B _{b1}	B _{b2}	B ₃	.
			14	15

$$b_A * b_B + b_A$$

$$20 * 15 + 20 = 320$$

$$b_{rB} * (b_s - M + 1) + b_{rB}$$

$$20 * (15 - 10 + 1) + 20 = 140$$

1.) LRU: 320

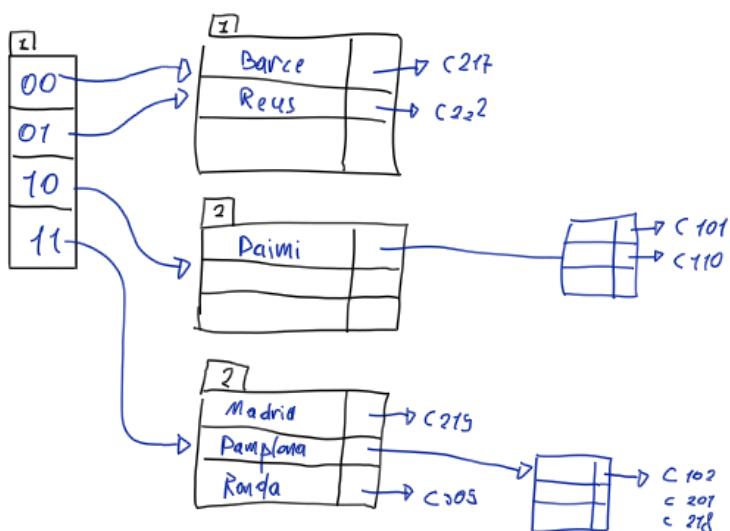
2.) MRU: 140

Tema 5 (10 p.). Ilustre la construcción de un índice Hash basado en la técnica de Asociación Extensible para la columna *nombre-sucursal* considerando para ello la tabla y los valore de hash indicados a continuación.

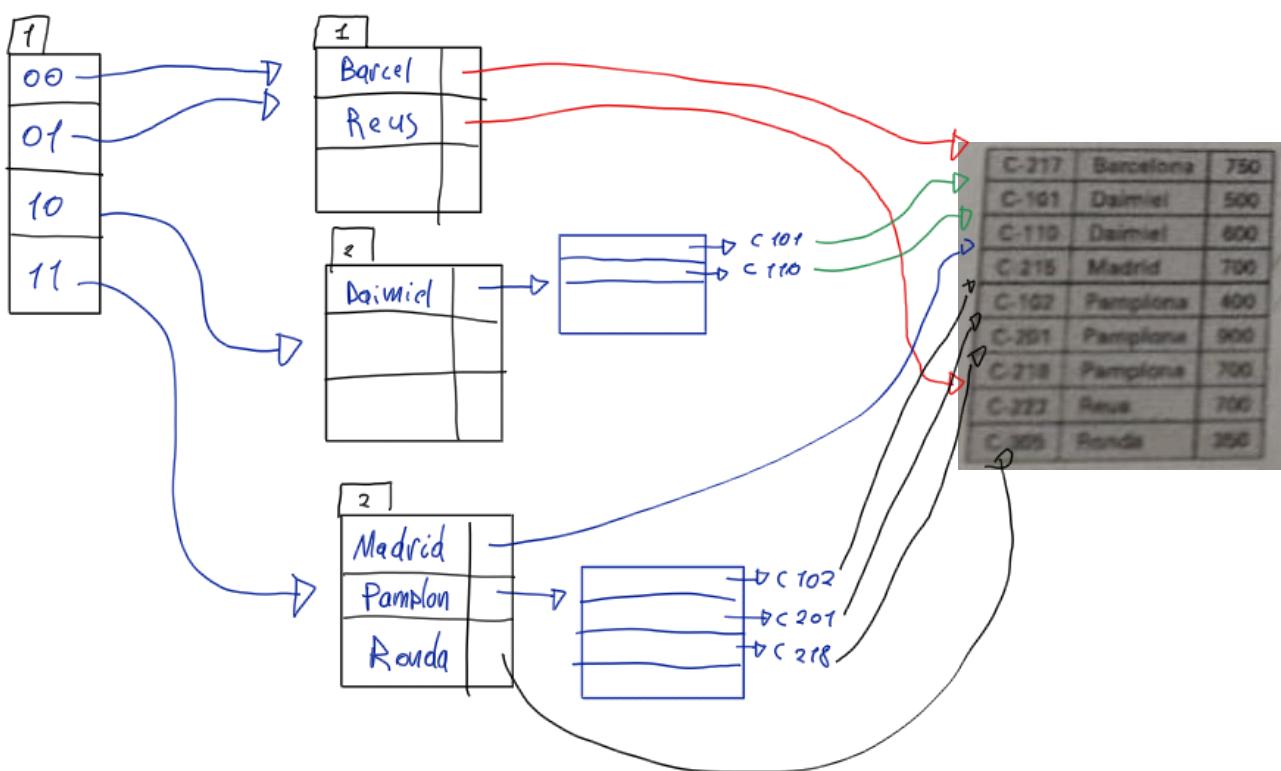
C-217	Barcelona	750
C-101	Daimiel	500
C-110	Daimiel	600
C-215	Madrid	700
C-102	Pamplona	400
C-201	Pamplona	900
C-218	Pamplona	700
C-223	Reus	700
C-309	Ronda	350

nombre-sucursal	binombre-sucursal
Barcelona	0010 1101 1111 1011 0010 1100 0011 0000
Daimiel	1010 0011 1010 0000 1100 0110 1001 1111
Madrid	1100 0111 1110 1101 1011 1111 0011 1010
Pamplona	1111 0001 0010 0100 1001 0011 0110 1101
Reus	0011 0101 1010 0110 1100 1001 1110 1011
Ronda	1101 1000 0011 1111 1001 1100 0000 0001

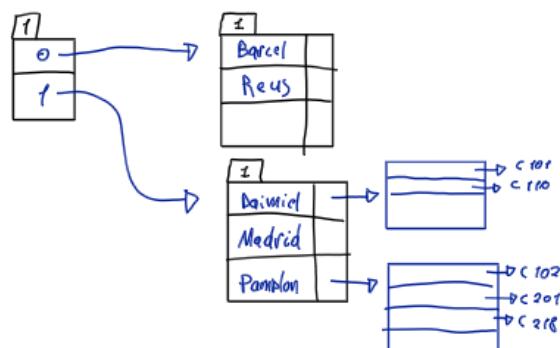
C-217	Barcelona	750
C-101	Daimiel	500
C-110	Daimiel	600
C-215	Madrid	700
C-102	Pamplona	400
C-201	Pamplona	900
C-218	Pamplona	700
C-223	Reus	700
C-309	Ronda	350



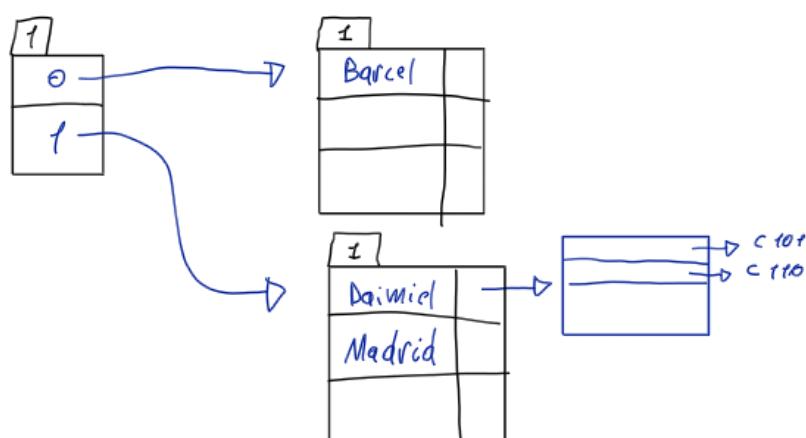
Paso 2



Paso 2



Paso 2



Paso 1



Tema 1 – 20 p. – Explique:

- a- Las medidas de rendimiento de los discos magnéticos. (5p)
- b- Los enfoques de organización de archivos un SGBD. (5p)
- c- Los tres tipos de índices estudiados y cuando conviene aplicarlos. (5p)
- d- Los pasos para el Procesamiento de Consultas. Haga y explique el diagrama. (5p.)

Tema 2 – 10 p. Conforme la siguiente tabla que corresponde al archivo de la relación EVALUACIÓN, en el que cada bloque del archivo corresponde a 1 fila, se pide:

- a- Construir un índice en forma de árbol B+ con nodos de 4 punteros para la clave primaria ID, suponiendo que los registros fueron ingresados (insertadas) según el orden de la columna NOMBRE.
- b- Construir un índice hash estático cerrado con cajones de 3 registros índices cuya función de asociación es “ $x \bmod 5$ ” sobre la columna PUNTAJE, siendo x el valor de cada fila en dicha columna.
- c- Explique para cada caso (a y b) anterior, si los registros índices apuntan o no a la tabla y por qué.

ID	NOMBRE	PUNTAJE
1	Preston Schwartz	10
2	Cathleen Steele	3
3	Tatyana Russo	11
4	Libby Madden	7
5	Orla Reid	8
6	Vivian Cherry	12
7	Kirk Jensen	5
8	Amanda Macias	1
9	Barry Morris	2
10	Lee Lopez	6
11	Elliott Fowler	4
12	Paula Johns	9

mod 5

3
0
1
0
2
1
0
0
4
3
2
0

Tema 3 – 5 p. Dadas las relaciones $r1(A, B, C)$ y $r2(C, D, E)$ con las siguientes propiedades: $r1$ tiene 30.000 tuplas, $r2$ tiene 90.000 tuplas, 100 tuplas de $r1$ caben en un bloque y 45 tuplas de $r2$ que caben en un bloque. Estímese el número de accesos a bloques requeridos utilizando las siguientes estrategias para la reunión “ $r1 \text{ join } r2 \text{ on } (r1.C = r2.C)$ ”:

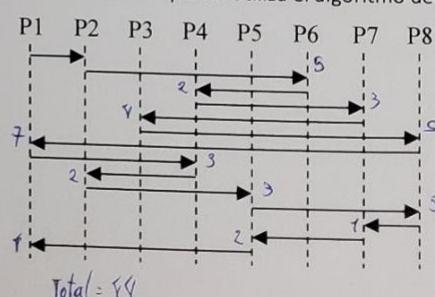
- a- Reunión en bucle anidado.
- b- Reunión en bucle anidado por bloques.
- c- Reunión por mezcla, suponiendo además que la relación $r1$ debe ser ordenada externamente disponiéndose de 3 bloques de memoria intermedia.
- d- Reunión en bucle anidado indexada, suponiendo que existe un índice primario disperso de árbol B+ en la columna C de la relación $r2$ con nodos de 60 punteros.

Tema 4 – 5 p. – Dada una operación de selección cualquiera basada en una condición de igualdad, explique en qué casos un SGBD utilizará los siguientes algoritmos e indique cual es el costo asintótico de los mismos:

- 1- Búsqueda Lineal.
- 2- Búsqueda Binaria.
- 3- Búsqueda en índice primario para un atributo clave.
- 4- Búsqueda en índice secundario para atributo no clave.

Tema 5 – 5 p extras. – La siguiente figura que muestra la planificación de lectura del disco en un SGBD que no utiliza el algoritmo del ascensor. Se pide:

- a- Explique el algoritmo del ascensor y el porqué de su utilidad.
- b- Muestre una planificación basada en el uso del algoritmo del ascensor suponiendo que:
 1. El recorrido del peine empieza en la pista 1,
 2. Se pueden atender varias solicitudes por parada.
 3. Las solicitudes llegan cada 1 ms.
 4. Las lecturas demoran 0,5 ms por solicitud.
 5. El peine se mueve a una razón de 2 pistas/ms
- c- Hallar la taza de atención de solicitudes por unidad de tiempo sin y con la aplicación del algoritmo del ascensor.



Tema 3 – 5 p. Dadas las relaciones $r1(A, B, C)$ y $r2(C, D, E)$ con las siguientes propiedades: $r1$ tiene 30.000 tuplas, $r2$ tiene 90.000 tuplas, 100 tuplas de $r1$ caben en un bloque y 45 tuplas de $r2$ que caben en un bloque. Estímese el número de accesos a bloques requeridos utilizando las siguientes estrategias para la reunión “ $r1 \text{ join } r2 \text{ on } (r1.C = r2.C)$ ”:

- a- Reunión en bucle anidado.
- b- Reunión en bucle anidado por bloques.
- c- Reunión por mezcla, suponiendo además que la relación $r1$ debe ser ordenada externamente disponiéndose de 3 bloques de memoria intermedia.
- d- Reunión en bucle anidado indexada, suponiendo que existe un índice primario disperso de árbol B+ en la columna C de la relación $r2$ con nodos de 60 punteros.

d-) $\frac{30.000}{100} \times C + b_{r1}$ $\frac{300}{45}$

$$V(A, r) = \frac{90.000}{45} = 2000$$

$$\lceil \log_{\frac{9}{2}} (V(A, r)) \rceil$$

$$30.000 \times \lceil \log_3 (2000) \rceil + 300 = 90.300$$

2do Parcial

Tema 5-10 p. - Dadas las relaciones A(a1, a2, ..., a20), B(b1, b2, ..., b12) y C(c1, c2, ..., c15), y la siguiente consulta:

```
select distinct A.a1, C.c1 from A join B (A.a2 = B.b3)
join C (C.c2 = B.b4) where A.a1 > 10 and B.b1 = 50;
```

Muestre:

1. Su traducción directa al álgebra relacional.
2. Los pasos para llegar a la expresión equivalente que minimice la cantidad de datos procesado por cada operación.
3. El árbol de evaluación de la expresión final.

* operación de distribución
* distribución de equivalencia conjuntiva

1. $\Pi_{A.a1, C.c1} \left(\sigma_{A.a1 > 10 \wedge B.b1 = 50} \left((A \bowtie_{A.a2 = B.b3} B) \bowtie_{C.c2 = B.b4} C \right) \right)$

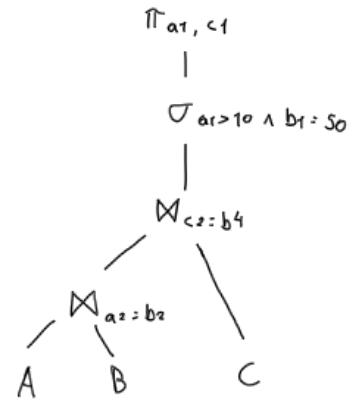
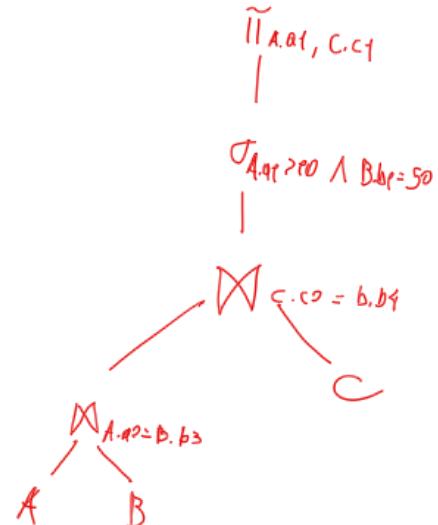
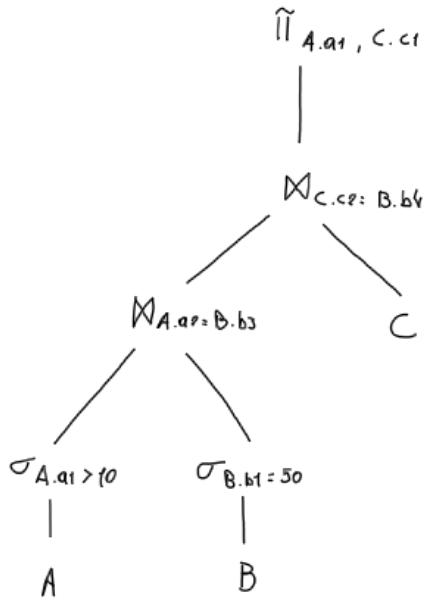
2. Regla 7 $\sigma_{\theta_1} (X \bowtie Y) = (\sigma_{\theta_1} X) \bowtie Y$

$$\Pi_{A.a1, C.c1} \left(\left(\sigma_{A.a1 > 10 \wedge B.b1 = 50} (A \bowtie_{A.a2 = B.b3} B) \right) \bowtie_{C.c2 = B.b4} C \right)$$

por Regla 7 selección distributiva en join

$$\Pi_{A.a1, C.c1} \left(\left(\left(\sigma_{A.a1 > 10} A \right) \bowtie_{A.a2 = B.b3} \left(\sigma_{B.b1 = 50} B \right) \right) \bowtie_{C.c2 = B.b4} C \right)$$

3.



Tema 6 - 5p – Determine si las siguientes planificaciones son o no secuenciables en cuanto a conflictos e indique cual es el orden de las transacciones en caso afirmativo.

T1	T2
Leer(A)	Escribir(A)
Escribir(A)	Escribir(A)

T1	T2
Leer(A)	Escribir(B)
Leer(B)	Escribir(A)

T1	T2	T3
Escribir(A)	Escribir(B)	Escribir(C)
	Leer(C)	

No
Sí
No.

No!

Si! T1, T3, T2

T1	T2	T3
Escribir(A)		Escribir(C)
	Escribir(B)	
		Leer(C)

Tema 4 (10 p.). Dada la consulta de abajo, proporcione una traducción inicial de la misma en álgebra relacional y luego proceda a ilustrar con la misma al menos dos casos de optimización conforme las reglas de equivalencias estudiadas.

select e.LNAME from EMPLEADO e
join TRABAJA_EN te on (te.EMPLEADO = e.ID)
join PROYECTO p on (p.ID = te.PROYECTO)
where p.FNOMBRE = 'AQUARIUS' and e.FECHA_INICIO >= '2000-01-01'

$$\Pi_{e.LName} \left(\sigma_{\text{O1} \wedge \text{O2}} \left([e \bowtie_{te.emp = e.ID} te] \bowtie_{p.ID = te.Proy} P \right) \right)$$

Regla 5 - Join commutativo

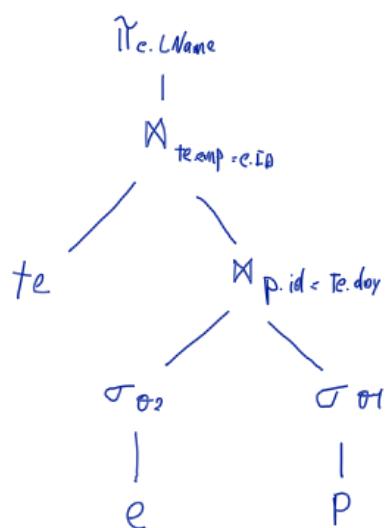
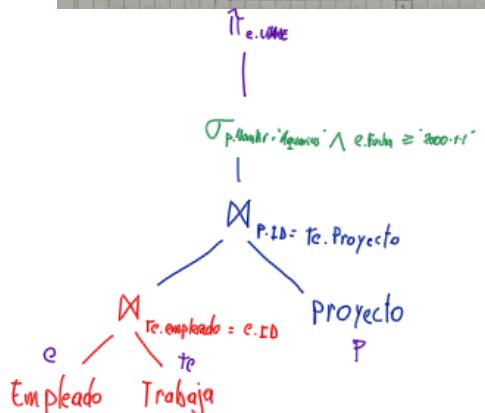
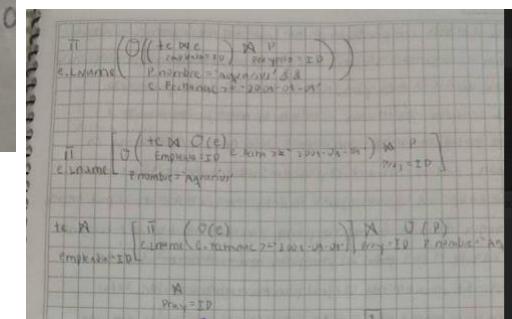
$$\Pi_{e.LName} \left(\sigma_{\text{O1} \wedge \text{O2}} \left([te \bowtie_{te.emp = e.ID} e] \bowtie_{p.ID = te.Proy} P \right) \right)$$

Regla 6 - Join asociativo

$$\Pi_{e.LName} \left(\sigma_{\text{O1} \wedge \text{O2}} \left(te \bowtie_{te.emp = e.ID} [e \bowtie_{p.ID = te.Proy} P] \right) \right)$$

Regla 7 - selección distributiva en join

$$\Pi_{e.LName} \left(\left(te \bowtie_{te.emp = e.ID} \left[(\sigma_{\text{O2}} e) \bowtie_{p.ID = te.Proy} P \right] \right) \right)$$



Tema 5 (10 p.). Explique en el contexto de los Protocolos del Control de Conurrencia que se entiende por: (1) Planificación, (2) Planificación Secuenciable, (3) Secuencialidad en Cuanto a Conflicto y (4) Secuencialidad en Cuanto a Vistas.

Protocolo de control de concurrencias

1.) Planificación: Es un ordenamiento de todas las operaciones de un conjunto de transacciones, conservando el orden interno

2.) Planificación Secuenciable: Una planificación de un conjunto de n transacciones es secuenciable si es equivalente a una planificación secuencial de estas n transacciones. La equivalencia puede ser en cuanto a conflictos o en cuanto a vistas.

3.) Secuencialidad en cuanto a Conflictos: Una planificación es secuenciable en cuanto a conflictos si son equivalentes en cuanto a conflictos a una planificación secuencial, esto es, si realizan operaciones sobre el mismo dato y al menos una de esas operaciones es escribir, el orden de esas operaciones debe mantenerse.

4.) Secuencialidad en cuanto a vistas: Una planificación es secuenciable en cuanto a vistas si son equivalentes en cuanto a vistas a una planificación secuencial, esto es, cumple los criterios:

- todas las operaciones que leen un valor inicial de un dato deben seguir leyendo el mismo valor
- todas las operaciones que leen un valor producido por otra operación deben seguir leyendo el mismo valor
- Si hay una operación escribir al final de una T, esta debe mantenerse al final de la T

Tema 6 (10 p.). Explique detalladamente las propiedades de las transacciones que deben ser aseguradas por un SGBD conforme al estándar SQL.

Propiedades de Transacciones

- | | |
|---------------|--|
| A Atomicity | Atomicidad: se ejecuta todo o nada |
| C Consistency | Consistencia: después de ejecutarse la BD debe ser consistente |
| I Isolation | Aislamiento: |
| D Durability | Persistencia: |

Tema 9 (10 p.). Desarrolle el algoritmo de Programación Dinámica para la optimización del orden de reunión en consultas. Explique detalladamente cual es el principio detrás del mismo, así como la complejidad computacional espacial y temporal que corresponde al mismo.

Tema 10 (10 p.). Ejemplifique la estrategia de la Semi Reunión para el procesamiento distribuido de reuniones. Explique al menos una ventaja de la misma que no se refiera a la transmisión de datos.

Bases de Datos II - Segundo Examen Parcial – 19/11/2018 – Duración 90 min.

Nombre: _____

✓ Tema 1 – 10 p. – Explique:

- a) Las formas de Almacenamiento Distribuido en BDD y sus ventajas.
- b) Que se entiende por Alta Disponibilidad y Reconfiguración en un Sistema de BDD.
- c) El Algoritmo del Luchador.

✓ Tema 2 – 10 p. – Muestre los pasos para la reunión de las siguientes relaciones mediante la estrategia de la semireunión. Luego, plantee el ahorro que conlleva el mismo en transmisión respecto a la estrategia más simple que realiza el menor movimiento de datos entre los sitios. Suponga que la consulta ha sido recibida en el Sitio 1 y que el resultado debe producirse en este mismo sitio.

R join S on (R.A2 = S.A2)

R@S1		S@S2		
A1	A2	A2	A3	A4
1	3	3	13	16
1	4	3	14	16
1	6	7	13	17
2	3	10	14	16
2	6	10	15	17
3	7	11	15	16
3	8	11	15	16
3	9	12	15	16

✓ Tema 3 – 10 p. – Explique: (i) el protocolo de control de bloqueo distribuido de Quórum de Consenso, (ii) que implica el que se cumpla la condición del protocolo y (iii) como pueden emularse los protocolos de Mayoría y Sesgado con el mismo.

✓ Tema 4 – 10 p. – Esquematice el protocolo de compromiso de 2 fases e indique como se procede: (a) en caso de falla del coordinador y (b) en caso de falla de un sitio participante.

✓ Tema 5 – 5 p. – Ilustre y describa la arquitectura de implementación de un Dataware.

~~Tema 6 – 5 p. – Explique las 3 V's que caracterizan a las iniciativas de Big Data. No Entra~~

Tema 2 – 10 p. – Muestre los pasos para la reunión de las siguientes relaciones mediante la estrategia de la semireunión. Luego, plantee el ahorro que conlleva el mismo en transmisión respecto a la estrategia más simple que realiza el menor movimiento de datos entre los sitios. Suponga que la consulta ha sido recibida en el Sitio 1 y que el resultado debe producirse en este mismo sitio.

R join S on (R.A2 = S.A2)

R@S1		S@S2		
A1	A2	A2	A3	A4
1	3	3	13	16
1	4	3	14	16
1	6	7	13	17
2	3	10	14	16
2	6	10	15	17
3	7	11	15	16
3	8	11	15	16
3	9	12	15	16

- Una estrategia posible para evitar esto es la siguiente:
 1. Calcular $\text{temp}_1 = \Pi_{R1 \cap R2}(r_1)$ en S_1 .
 2. Enviar temp_1 a S_2 .
 3. Calcular $\text{temp}_2 = r_2 <\text{join}> \text{temp}_1$ en S_2 .
 4. Enviar temp_2 a S_1 .
 5. Calcular $r_1 <\text{join}> \text{temp}_2$ en S_1 .
El resultado es el mismo que $r_1 <\text{join}> r_2$.

Paso 1: $\text{temp}_1 = \Pi_{R1 \cap R2}(r_1)$ en $S_1 \rightarrow \text{temp}_1 = \begin{bmatrix} 3 \\ 4 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix}$

Sitio 1 – R	
A1	A2
1	3
1	4
1	6
2	3
2	6
3	7
3	8
3	9

Sitio 2 – S		
A2	A3	A4
3	13	16
3	14	16
7	13	17
10	14	16
10	15	17
11	15	16
11	15	16
12	15	16

$\rho \times 3$

* Π elimina los valores duplicados

1. $\text{temp}_1 = \Pi_{R \cap S}(R)$ en $S_1 = \begin{bmatrix} 3 \\ 4 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix}$ 0 udt

2. $\text{temp}_1 \rightarrow S_1 \rightarrow S_2$ 6 Updates transmitidos

$\Theta = \text{igualdad}(R \cap S)$

$R.Q_1 = S.Q_2$

3. $\text{temp}_2: S \times_{\Theta} \text{temp}_1 = \begin{bmatrix} 3 & 13 & 16 \\ 3 & 14 & 16 \\ 7 & 13 & 17 \end{bmatrix}$ 0 udt

X es como un left Join/
creo

4. $\text{temp}_2 \rightarrow S_2 \rightarrow S_1$ 9 udt

5. Resultado: $R \bowtie \text{temp}_2$ total 15 udt

$S \rightarrow S_2 \rightarrow S_1 \rightarrow 24$ udt

$S \bowtie R$

Tema 1 – 10 p. – Explique:

- a- Las formas de Almacenamiento Distribuido en BDD y sus ventajas.
- b- Que se entiende por Alta Disponibilidad y Reconfiguración en un Sistema de BDD.
- c- El Algoritmo del Luchador.

Tema 4 – 10 p. – Esquematicice el protocolo de compromiso de 2 fases e indique como se procede: (a) en caso de falla del coordinador y (b) en caso de falla de un sitio participante.

TEMA 4:

① FASE 1:

- Comienza cuando Ci recibe "ultimo preparado" de T.
- Si el coordinador en su bitácora <Preparar T>, y fija su grabación.
- Envió a todos los sitios Sis el mensaje <Preparar T>.
- En los Sis, el gerador decide si confirmar o no su parte de T:
 - Si decide que no <Abortar T> o su bitácora y envía el mensaje <Abortar T> al T.
 - Si decide que si & procede <Preparar T> a su bitácora, refresca su grabación y envía el mensaje <Preparar T> al T.

② FASE 2:

- Comienza cuando Ci recibe todo los respuestas de todos los Sis o cuando transcurrido un tiempo no recibe todo.
- Si recibe todos los respuestas y todos son <Preparado T>, Ci confirma T y envía el mensaje <Confirmado T> a los Sis.
 - Si se cil recibir el mensaje, se pone el registro <Confirmado T> a su bitácora y fija su grabación.
 - Si no recibe todos los respuestas o alguno es <Abortar T>, Ci abandona T y envía el mensaje <Abortado T> a los Sis.
 - Si se cil recibir el mensaje, se pone el registro <Abandonado T> a su bitácora y fija su grabación.

③ FALLA DEL COORDINADOR:

En este caso los Sis deben decidir el destino de T. Se fijan en su bitácora y:

- Si contiene <Confirmado T>, procede a confirmar T.
- Si contiene <Abandonado T>, procede a abandonar T.
- Si contiene <Preparado T> significa que no se confirmó y procede a abandonar T.

Si no contiene ninguno de los 3, solo puede decir que algunos contienen <Preparado T> y deben esperar que el coordinador recuperé para decidir el destino de T.

TEMA 4:

④ FALLO DE UN SITIO PARTICIPANTE:

Si un Sito falla, cierra, durante o después de una transacción T, una vez finalizado se fija en registro de su bitácora para decidir el destino de T.

- Si se contiene <Confirmado T>, realiza rebocer (T).
- Si se contiene <Abandonado T>, realiza deshacer (T).
- Si se contiene <Preparado T> coincide con la idea que hacer:
 - Si se responde <Confirmado T> se reboca (T).
 - Si se responde <Abandonado T> se realiza deshacer (T).
- Si se contiene resto en su bitácora significa que se fijó anteriormente de la transacción por lo que se ya abrió T y procede a realizar deshacer (T).

Fase 1: Comienza luego de la última sentencia de T

- Ci guarda <Preparar T> en bitácora y graba
- Ci envía <Preparar T> a todos

* Por cada sitio participante:

- Si NO puede añadir: guarda <No T> en bitácora
envía <Abortar T> a Ci
- Si puede añadir: guarda <Preparado T> a bitácora y graba
envía <Preparado T> a Ci

Fase 2: Comienza cuando Ci recibe las respuestas de todos

- o cuando pasa cierto tiempo

* Si recibe todos, ej todos son <Preparado T>:

- Ci guarda <Comprobado T> en bitácora y graba
- Ci envía <Comprobado T> a todos los sitios

* Por cada sitio: cuando recibe <Comprobado T>

- Compite su parte de T
- guarda <Comprobado T> a bitácora y graba

* Si no recibe todas las respuestas, o una es <Abandonado T>

- Ci guarda <Abandonado T> en bitácora y graba
- Ci envía <Abortar T> a todos los sitios

* Por cada sitio: cuando recibe <Abortar T>

- Aborta su parte de T y ejecuta recuperación
- guarda <Abandonado T> a bitácora y graba

a) En caso de falla de controlador

- los sitios deben decidir el destino de T
- si uno tiene <Comprobado T>, T debe ser comprobada
- .. <Abortar T> .. abortada
- Si ninguno tiene <Preparado T>, Ci no decidió confirmar T, se aborta
- Si Nda. entonces alguno esta prepa

(a) Defina el concepto de Transacción en SGBDs. (b) Indique y defina las fases del Ciclo de Vida de una Transacción. (b) Indique y defina las propiedades que un SGBD debe asegurar para las Transacciones.

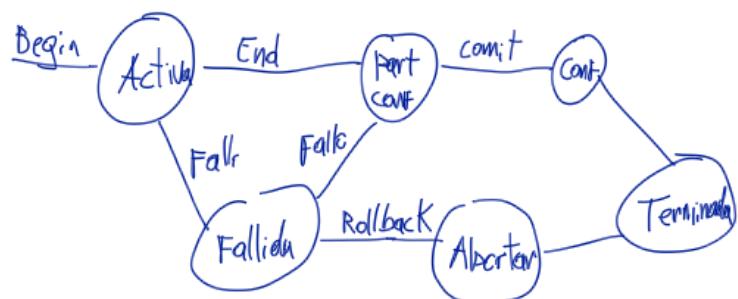
a.) Un conjunto de operaciones . . . leer/escibir - - - una unidad

b.) Activa al comenzar

Parcialmente Confirmada aq se realizan controles de concurrencia, luego de End
Confirmada confirmación y concluye la ejecución
Fallida

Abortada se deshacen los efectos

Terminada cuando desaparece



T1
leer B leer C leer A

A - Determine si las siguientes planificaciones son o no secuenciables en cuanto a conflictos, e indique cual es el orden de las transacciones en caso afirmativo.

P1	
T1	T2
Leer(A)	Escribir (A)
Escribir(A)	Escribir(A)

No

P2	
T1	T2
Leer(A)	Escribir (B)
Leer(B)	Escribir (A)

No

P3		
T1	T2	T3
Escribir(A)	Escribir(B)	Escribir(C)
Leer(C)		Escribir(C)

Yes

T1	T2	T3
Esc A		Esc C
Esc B	Leer C	

en cuanto a conflictos
si hay escribir, mantener orden

orden: T1, T3, T2

B - Determine si las siguientes planificaciones son o no secuenciables en cuanto a vistas, e indique cual es el orden de las transacciones en caso afirmativo.

P1	
T1	T2
Leer(A)	Escribir (A)
Escribir(A)	Escribir(A)

Si

P2

T1	T2
Leer(A)	Escribir (B)
Leer(B)	Escribir (A)

No

T1, T2

P3		
T1	T2	T3
Leer(A)	Leer(B)	Escribir(B)
Escribir(A)	Escribir(B)	Leer(A)

T1	T2	T3
Leer A	Leer B	Escribir B
ESC A	ESC B	Leer A
		ESC B

Si :

T1, T2, T3

- si leyó valor inicial, mantener
- " valor modificado, "
- " escribió al final de una T. mantener al final de esa T"

1^{er} Final

Tema 1 (10 p.). Explique en qué consisten los problemas que deben evitarse en la detección de Interbloqueos en un Sistema de Base de Datos Distribuida.

Tema 2 (10 p.).

- Desarrolle los pasos correspondientes para la semi-reunión de las relaciones ilustradas abajo. Suponga que la consulta ha sido recibida en el Sitio 2.
- Explique o fundamente al menos dos (2) ventajas de este enfoque.

R join S on (R.A2 = S.A2)

R@S1		S@S2		
A1	A2	A2	A3	A4
1	3	3	13	16
1	4	3	14	16
1	6	7	13	17
2	3	10	14	16
2	6	10	15	17
3	7	11	15	16
3	8	11	15	16
3	9	12	15	16

Tema 3 (10 p.). Explique los tipos de índices estudiados conforme su organización/implementación física y para qué tipo de consultas resultan apropiados.

Tema 4 (10 p.). Explique: (i) el protocolo de control de bloqueo distribuido de Quórum de Consenso, (ii) que implica el que se cumpla la(s) condición(es) del protocolo y (iii) la configuración de parámetros que deben aplicarse para emular los protocolos de Mayoría y Sesgado.

Tema 5 (10 p.). El ciclo para el Procesamiento de Consultas en un SGBD. Haga el diagrama y explique cada paso.

Tema 6 (10 p.). Dadas las relaciones A(a1, a2, ..., a20), B(b1, b2, ..., b12) y C(c1, c2, ..., c15), y la siguiente consulta:

```
select distinct A.a1, C.c1 from A join B on (A.a2 = B.b3)
join C on(C.c2 = B.b4) where A.a1 > 10 and B.b1 = 50;
```

Muestre:

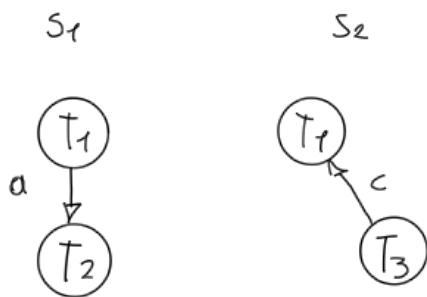
- Su traducción directa al álgebra relacional.
- Detalle al menos 3 optimizaciones basadas en expresiones equivalentes que pueden aplicarse. Debe indicar la regla de equivalencia aplicada.
- El árbol de evaluación de la expresión final, conforme el punto anterior

Tema 7 (10 p.)

- Ilustre y describa la arquitectura de implementación de un Dataware.
- Explique las 3 V's que caracterizan a las iniciativas de Big Data.

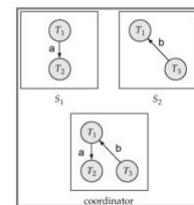
Tema 8 (10 p.) Detalle apropiadamente como se implementa la organización física de bloques denominada *estructura de páginas por ranuras* para el almacenamiento de registros.

Tema 1 (10 p.). Explique en qué consisten los problemas que deben evitarse en la detección de Interbloqueos en un Sistema de Base de Datos Distribuida.



Se construye un grafo de espera global

Detección de Ciclos: Ciclos Falsos

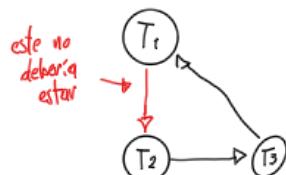


- Luego ocurre en el sgte. orden que:
1. T2 libera el dato a
 2. S1 envía el mensaje de eliminar el arco $T_1 \rightarrow T_2$
 3. T2 solicita un bloqueo sobre c en el sitio S2, pero T3 tiene bloqueado el dato c
 4. S2 envía el mensaje de añadir el arco $T_2 \rightarrow T_3$

Ciclos falsos

Si T_2 libero' a pero no llego' el mensaje de S_1 de eliminar el arco $T_1 \rightarrow T_2$

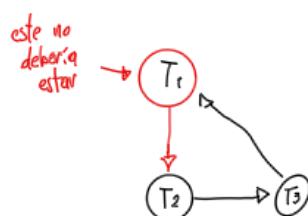
Si el mensaje de S_1 llega luego de bloquear $T_2 \rightarrow T_3$



Retroceso no relacionado

Cuando una transacción falla, se elimina su arco, pero el coordinador no lo detecta y tb recibe otra solicitud de bloqueo
El coordinador retrocederá una de las transacciones innecesariamente

$(T_2 \circ T_3)$



Tema 2 (10 p.).

- Desarrolle los pasos correspondientes para la semi-reunión de las relaciones ilustradas abajo. Suponga que la consulta ha sido recibida en el Sitio 2.
- Explique o fundamente al menos dos (2) ventajas de este enfoque.

R join S on (R.A2 = S.A2)

R@S1		S@S2		
A1	A2	A2	A3	A4
1	3	{3	13	16
1	4	3	14	16
1	6	7	13	17
2	3	{10	14	16
2	6	10	15	17
3	7	{11	15	16
3	8	11	15	16
3	9	12	15	16

a-)

$$1. \text{ temp}_1 = \prod_{R.A2=S.A2} (S) = \begin{bmatrix} 3 \\ 7 \\ 10 \\ 11 \\ 12 \end{bmatrix} \quad 0 \text{ udt}$$

$$2. \text{ Temp}_1 \rightarrow S_2 \rightarrow S_r \quad 5 \text{ udt}$$

$$3. \text{ Temp}_2 = R \times \text{temp}_1 = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 3 & 7 \end{bmatrix} \quad 0 \text{ udt}$$

$$4. \text{ Temp}_2 \rightarrow S_r \rightarrow S_2 \quad 6 \text{ udt}$$

$$5. \text{ Resultado} = S \times \text{temp}_2 = \begin{bmatrix} 1 & 3 & 13 & 16 \\ 1 & 3 & 14 & 16 \\ 2 & 3 & 13 & 16 \\ 2 & 3 & 14 & 16 \\ 3 & 7 & 13 & 17 \end{bmatrix}$$

b-) Dos ventajas de este enfoque

* Reducción de costos de transmisión datos

\prod udt vs mandar R q S₂ = 16 udt

* La carga de trabajo se distribuye entre los sitios, se aprovechan mejor los recursos de la red

?

Tema 3 (10 p.). Explique los tipos de índices estudiados conforme su organización/implementación física y para qué tipo de consultas resultan apropiados.

Índices Ordenados: En los cuales los valores se guardan en forma secuencial en relación a los datos de la tabla. Son apropiados para búsquedas en rango

Índices Asociativos: En los cuales los valores están distribuidos en base a (Hash) una función hash. Son apropiados para búsquedas en igualdad

Índice Bitmap: Almacena un array de bits que contiene tantas columnas como filas de la tabla, y tantas filas como valores discretos haya, Es ideal para tablas con valores discretos (ej: género: {h, m}) Es bueno para búsquedas con and/or en búsquedas compuesta

Tema 4 (10 p.). Explique: (i) el protocolo de control de bloqueo distribuido de Quórum de Consenso, (ii) que implica el que se cumpla la(s) condición(es) del protocolo y (iii) la configuración de parámetros que deben aplicarse para emular los protocolos de Mayoría y Sesgado.

(1) Es un protocolo para poder conceder bloqueos de 1/e a sitios. Se asignan pesos a cada sitio, siendo s la suma de todos los pesos y se define Q_r y Q_w . con las condiciones $Q_r + Q_w > s$ y $2Q_w > s$

* Si un sitio desea realizar una lectura, debe bloquear tantos sitios tal que la suma de los pesos sea $\geq Q_r$

* Lo mismo cuando un sitio desea escribir con Q_w

(2) $Q_r + Q_w > s$ si se hace una lectura, no se puede hacer a la vez una escritura y vice-versa

$2Q_w > s$ solo un sitio puede escribir un dato a la vez

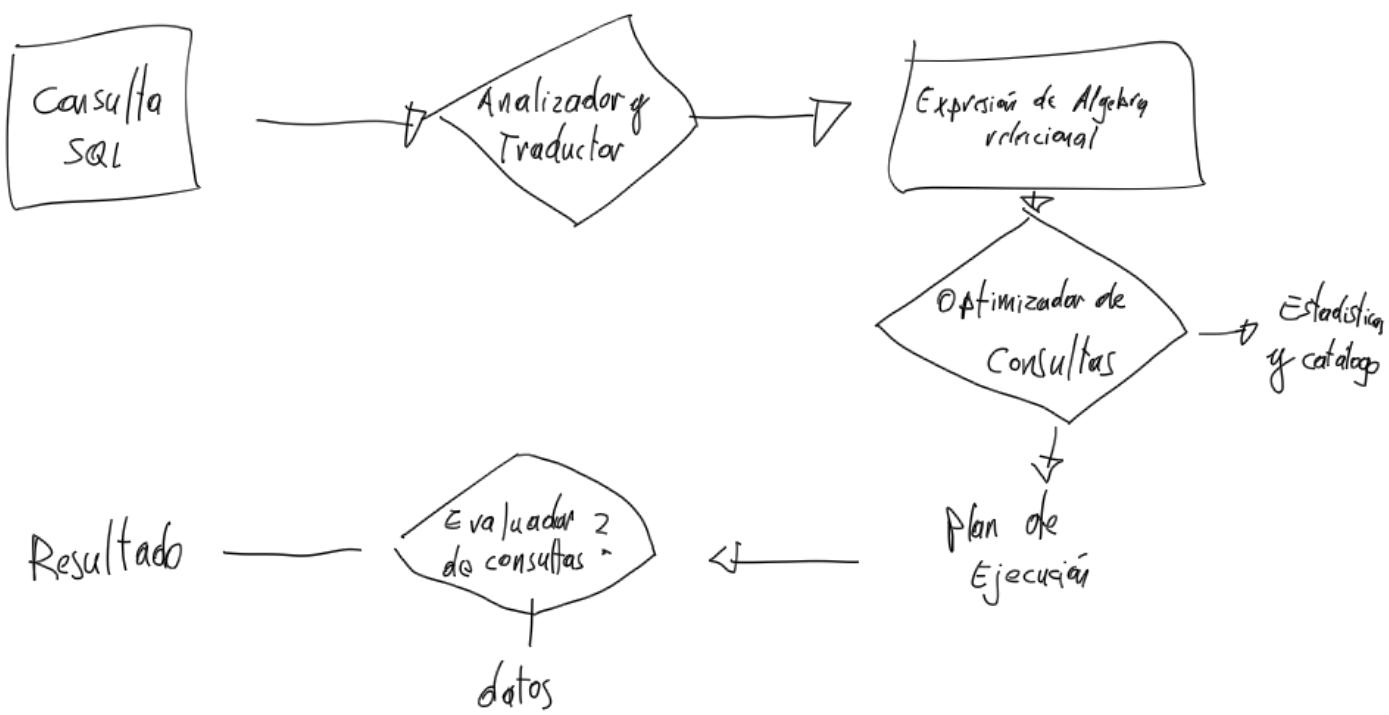
3. Protocolo de Mayoría \checkmark o N? N = N° Sitios

$Q_r = Q_w = S/2 + 1$ por lo tanto para cada operación se requerirá de aprobación de la mayoría de los sitios

Protocolo Sesgado

$Q_r = 1$ $Q_w = S$ por lo tanto la escritura es completamente bloqueante

Tema 5 (10 p.). El ciclo para el Procesamiento de Consultas en un SGBD. Haga el diagrama y explique cada paso.



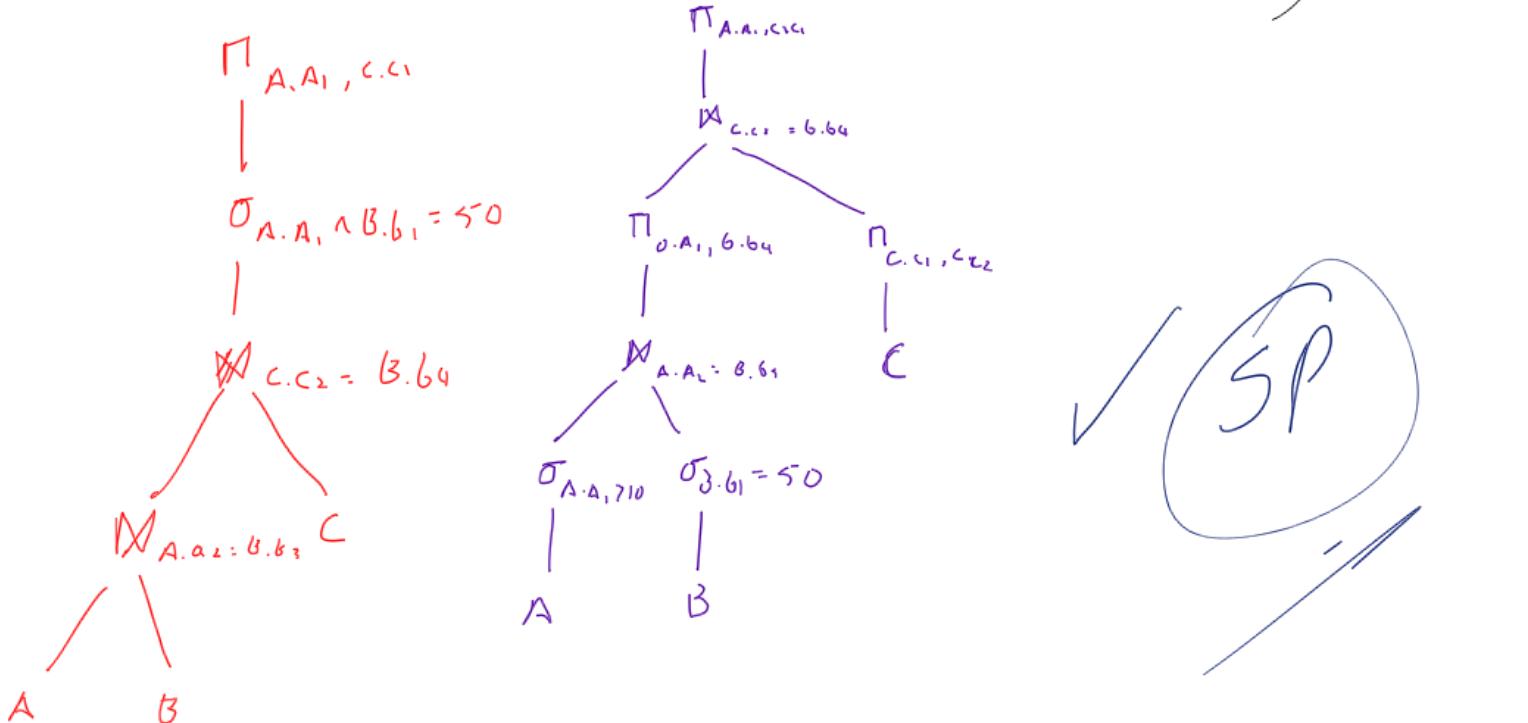
Tema 6 (10 p.). Dadas las relaciones A(a1, a2, ..., a20), B(b1, b2, ..., b12) y C(c1, c2, ..., c15), y la siguiente consulta:

```
select distinct A.a1, C.c1 from A join B on (A.a2 = B.b3)
join C on(C.c2 = B.b4) where A.a1 > 10 and B.b1 = 50;
```

Muestre:

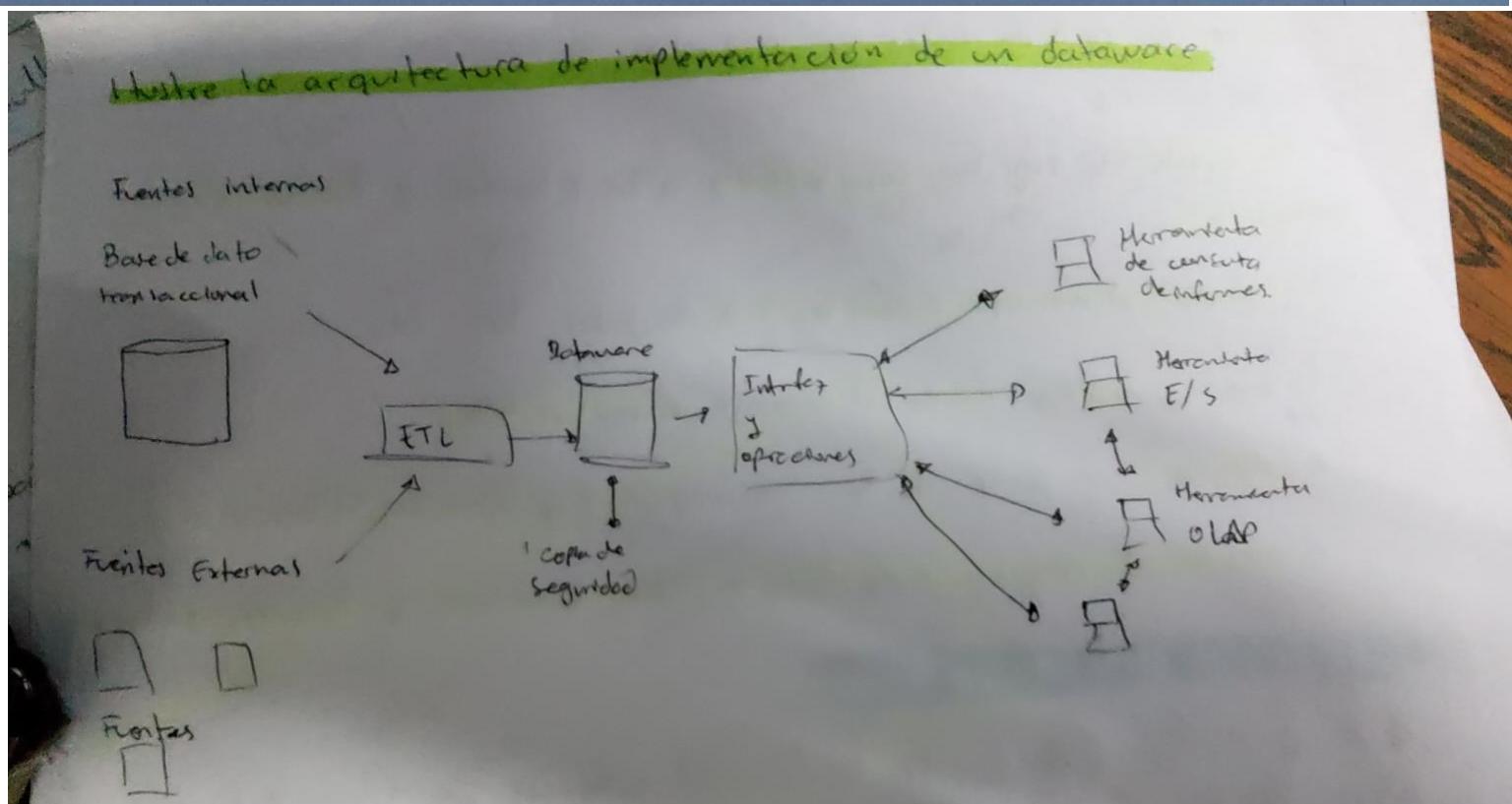
1. Su traducción directa al álgebra relacional.
2. Detalle al menos 3 optimizaciones basadas en expresiones equivalentes que pueden aplicarse. Debe indicar la regla de equivalencia aplicada.
3. El árbol de evaluación de la expresión final, conforme el punto anterior

$$\begin{aligned} \overline{\Pi}_{A.a1, C.c1} & \left(\sigma_{A.a1 > 10 \wedge B.b1 = 50} \left[\begin{array}{c} A \bowtie_{A.a2 = B.b3} (B \bowtie_{C.c2 = B.b4} C) \end{array} \right] \right) \\ \overline{\Pi}_{A.a1, C.c1} & \left(\sigma_{A.a1 > 10 \wedge B.b1 = 50} \left[(A \bowtie_{A.a2 = B.b3} B) \bowtie_{C.c2 = B.b4} C \right] \right) \quad 1^{\circ} \quad \text{selección asociativa} \\ \overline{\Pi}_{A.a1, C.c1} & \left(\left[\sigma_{A.a1 > 10 \wedge B.b1 = 50} (A \bowtie_{A.a2 = B.b3} B) \right] \bowtie_{C.c2 = B.b4} C \right) \quad 2^{\circ} \quad \text{selección distributiva} \\ \overline{\Pi}_{A.a1, C.c1} & \left(\left[\left(\sigma_{A.a1 > 10} (A) \bowtie_{A.a2 = B.b3} \sigma_{B.b1 = 50} (B) \right) \bowtie_{C.c2 = B.b4} C \right] \right) \quad 3^{\circ} \quad \text{selección distributiva} \\ \overline{\Pi}_{A.a1, C.c1} & \left(\overline{\Pi}_{A.a1, B.b1} \left(\left(\sigma_{A.a1 > 10} (A) \bowtie_{A.a2 = B.b3} \sigma_{B.b1 = 50} (B) \right) \bowtie_{C.c2 = B.b4} \overline{\Pi}_{C.c1, C.c2} (C) \right) \right) \quad \text{proyección distributiva} \end{aligned}$$

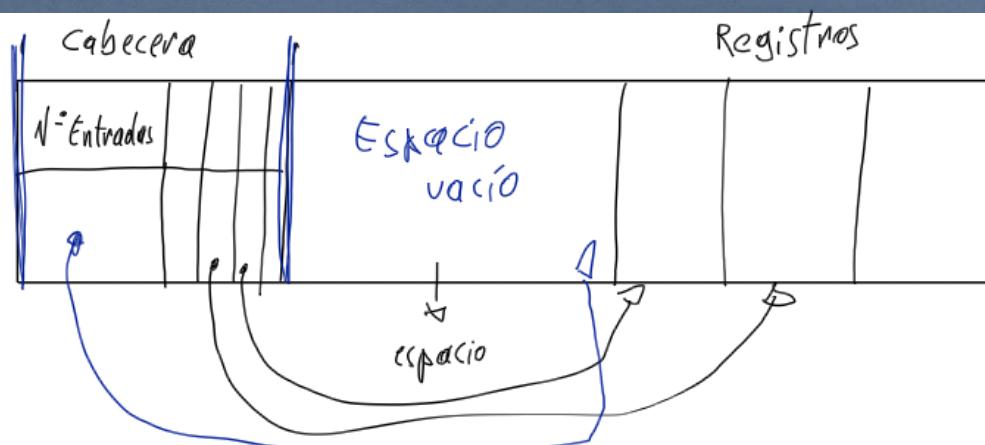


Tema 7 (10 p.)

- Ilustre y describa la arquitectura de implementación de un Dataware.
- Explique las 3 V's que caracterizan a las iniciativas de Big Data.



Tema 8 (10 p.) Detalle apropiadamente como se implementa la organización física de bloques denominada estructura de páginas por ranuras para el almacenamiento de registros.



Explique detalladamente el protocolo de Control de Conurrencias de dos fases
este protocolo exige que cada transacción realice sus
peticiones en dos fases

Fase de crecimiento: la transacción puede obtener bloqueos
pero no liberarlos

Fase de decrecimiento: puede liberar bloqueos pero no obtener
nuevos.

Explique además las variantes del mismo

Bloqueo estricto en dos fases.

los bloques exclusivos se liberan al final de la transacción
- se evitan los retrocesos en cascada

Bloqueo riguroso en dos fases:

todos los bloques se liberan al final ...
- las transacciones se pueden secuenciar en el
orden en que se comprometen.

Tema 2 - 10 p. - La siguiente tabla corresponde al archivo de la relación cliente, en el que cada bloque del archivo corresponde a 1 fila, se pide:

- a- Construir un índice ordenado en forma de árbol B+ con nodos de 5 punteros para la clave primaria id, suponiendo que los registros fueron insertados según el orden de los valores de la columna nombre.
- b- Construir un índice hash estático cerrado con cajones de 5 elementos cuya función de asociación es " $x \bmod 4$ " sobre la columna saldo, siendo x el valor de cada fila en dicha columna.
- c- Explique a donde apuntan los registros índices en cada caso y de qué tipo son los mismos teniendo en cuenta lo anterior.

ID	Nombre	Saldo
1	Griffith Berger	256
2	Philip Cherry	456
3	Kevin Conley	366
4	Tad Bird	164
5	Derek Bentley	166
6	Upton Lane	935
7	Branden Ortiz	313
8	Axel Shaw	213
9	Wayne Soto	544
10	Fuller Knowles	590
11	Merrill Bishop	159
12	Chadwick Hayden	657
13	David Jarvis	360
14	Jakeem Velez	203
15	Andrew Wolfe	953

orden

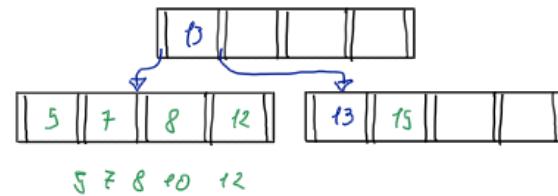
15	-
8	-
7	-
12	-
13	-
5	-
10	-
1	-
14	-
3	-
11	-
2	-
4	-
6	-
9	-

(a)

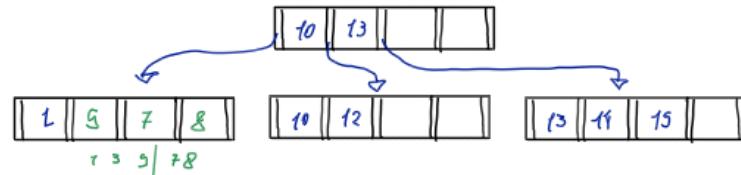
1º cargo 7, 8, 12, 15



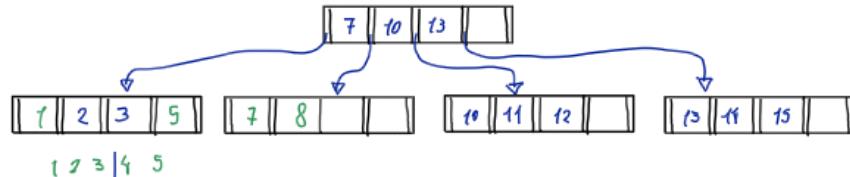
2º cargo 13, 5



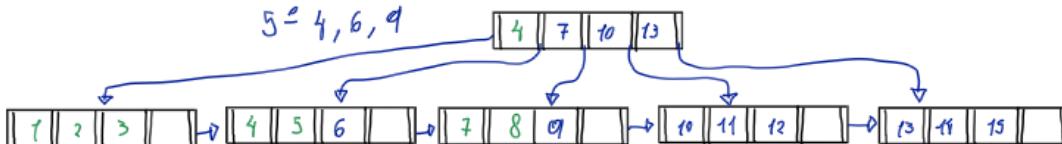
3º 10, 1, 14



4º 3, 11, 2,



5º 4, 6, 9



Tema 2 – 10 p. – La siguiente tabla corresponde al archivo de la relación cliente, en el que cada bloque del archivo corresponde a 1 fila, se pide:

- a- Construir un índice ordenado en forma de árbol B+ con nodos de 5 punteros para la clave primaria id, suponiendo que los registros fueron insertados según el orden de los valores de la columna nombre.
- b- Construir un índice hash estático cerrado con cajones de 5 elementos cuya función de asociación es " $x \bmod 4$ " sobre la columna saldo, siendo x el valor de cada fila en dicha columna.
- c- Explique a donde apuntan los registros índices en cada caso y de qué tipo son los mismos teniendo en cuenta lo anterior.

ID	Nombre	Saldo
1	Griffith Berger	256
2	Philip Cherry	456
3	Kevin Conley	366
4	Tad Bird	164
5	Derek Bentley	166
6	Upton Lane	935
7	Branden Ortiz	313
8	Axel Shaw	213
9	Wayne Soto	544
10	Fuller Knowles	590
11	Merrill Bishop	159
12	Chadwick Hayden	657
13	David Jarvis	360
14	Jakeem Velez	203
15	Andrew Wolfe	953

mod 4

0
0
2
0
2
3
1
1
0
2
3
1
0
3
1

0

256	
456	
164	
544	
360	

1

313	
213	
657	
953	

2

366	
166	
590	

3

935	
159	