

1. Sean tres transacciones $T_1 = \{O_{1,1}, O_{1,2}\}$, $T_2 = \{O_{2,1}, O_{2,2}, O_{2,3}\}$ y $T_3 = \{O_{3,1}, O_{3,2}\}$ donde $O_{i,j}$ es la operación j de la transacción i ($1 \leq i \leq 3$)

- a) **(0,2 puntos)** ¿Qué es un plan de ejecución de un conjunto de transacciones? Pon un ejemplo de ordenación de las operaciones de T_1 , T_2 y T_3 que no sea un plan, otro con un plan en serie y otro con un plan concurrente.

Un plan de ejecución de un conjunto de transacciones es un ordenamiento de las operaciones de las transacciones en el que las operaciones de cada transacción aparecen en el plan en el mismo orden que aparecen en la transacción.

$No_Plan = \{O_{1,2}, O_{2,1}, O_{2,2}, O_{2,3}, O_{1,1}, O_{3,1}, O_{3,2}\}$

$Plan_Serie = \{O_{1,1}, O_{1,2}, O_{2,1}, O_{2,2}, O_{2,3}, O_{3,1}, O_{3,2}\}$ (Hay seis planes en serie: $T_1 \rightarrow T_2 \rightarrow T_3$, $T_1 \rightarrow T_3 \rightarrow T_2$, $T_2 \rightarrow T_1 \rightarrow T_3$, $T_2 \rightarrow T_3 \rightarrow T_1$, $T_3 \rightarrow T_1 \rightarrow T_2$, $T_3 \rightarrow T_2 \rightarrow T_1$)

$Plan_concurrente = \{O_{1,1}, O_{2,1}, O_{2,2}, O_{1,2}, O_{2,3}, O_{3,1}, O_{3,2}\}$ (hay muchos planes posibles)

- b) **(0,15 puntos)** Comprueba si el protocolo **B2F Implícito** admite el plan P1 (es el mismo que en la cuestión anterior) dando una traza de su ejecución. Usa la tabla para contestar esta pregunta. **INDICA CLARAMENTE SI EL PLAN ES ACEPTADO O NO, en caso de no ser aceptado, explica el motivo.**

$P1 = r_3(z), r_1(x), w_1(x), r_1(y), c_1, r_2(x), r_3(x), c_3, w_2(x), w_2(y), c_2$

t	T ₁	T ₂	T ₃	Estado de bloqueo de los elementos de datos		
				x	y	z
				Desbloqueado	Desbloqueado	Desbloqueado
t ₀			r ₃ (z) BL Implícito	Desbloqueado	Desbloqueado	BL(T ₃)
t ₁	r ₁ (x) BL Implícito			BL(T ₁)	Desbloqueado	BL(T ₃)
t ₂	w ₁ (x) BE Implícito			BE(T ₁)	Desbloqueado	BL(T ₃)
t ₃	r ₁ (y) BL Implícito			BE(T ₁)	BL(T ₁)	BL(T ₃)
t ₄	c ₁ (Desbloqueo)			Desbloqueado	Desbloqueado	BL(T ₃)
t ₅		r ₂ (x) BL Implícito		BL(T ₂)	Desbloqueado	BL(T ₃)
t ₆			r ₃ (x) BL Implícito	BL(T ₂) y BL(T ₃)	Desbloqueado	BL(T ₃)
t ₇			c ₃ (Desbloqueo)	BL(T ₂)	Desbloqueado	Desbloqueado
t ₈		w ₂ (x) BE Implícito		BE(T ₂)	Desbloqueado	Desbloqueado
t ₉		w ₂ (x) BE Implícito		BE(T ₂)	BE(T ₂)	Desbloqueado
t ₁₀		c ₂ (Desbloqueo)		Desbloqueado	Desbloqueado	Desbloqueado

El plan es aceptado por el protocolo B2F Implícito

- c) **(0,15 puntos)** Pon un ejemplo de un plan con tres transacciones en el que, con el protocolo B2F explícito, se daría el problema del bloqueo mortal. Da una traza de su ejecución de tu plan que incluya los bloqueos y el punto en el que se produciría el bloqueo mortal. Usa la tabla para la ejecución (pueden sobrarte columnas en la parte de los estados de los elementos de datos). ¿Cómo se resolvería el problema?

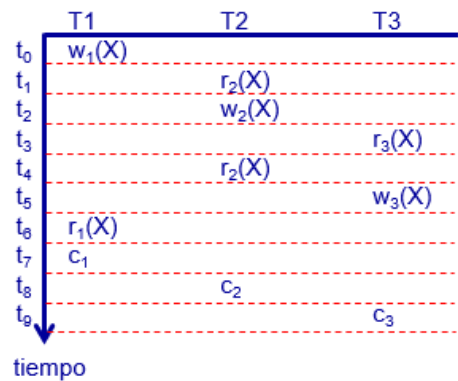
Plan = $\{r_3(y), r_1(x), r_2(x), w_1(x), w_2(x), c_3, c_1, c_2\}$

				Estado de bloqueo de los elementos de datos		
t	T ₁	T ₂	T ₃	x	y	
				Desbloqueado	Desbloqueado	
t ₀			BL(y)	Desbloqueado	BL(T ₃)	
t ₁			r ₃ (y)	Desbloqueado	BL(T ₃)	
t ₂	BL(x)			BL(T ₁)	BL(T ₁)	
t ₃	r ₁ (x)			BL(T ₁)	BL(T ₁)	
t ₄		BL(x)		BL(T ₁) y BL(T ₂)	BL(T ₁)	
t ₅		r ₂ (x)		BL(T ₁) y BL(T ₂)	BL(T ₁)	
t ₆	BE(x) espera			BL(T ₁) y BL(T ₂)		
t ₇		BE(x) espera		BL(T ₁) y BL(T ₂)		
t ₈						
t ₉						
t ₁₀						
t ₁₁						
t ₁₂						
t ₁₃						
t ₁₄						
t ₁₅						
t ₁₆						
t ₁₇						
t ₁₈						
t ₁₉						
t ₂₀						
t ₂₁						
t ₂₂						
t ₂₃						
t ₂₃						
t ₂₃						
t ₂₃						

Para resolver el bloqueo mortal el sistema abortará alguna de las transacciones implicadas (en este caso T1 o T2) y luego volvería a lanzarla.

2. (0,25 puntos)

Sea el siguiente plan de ejecución de tres transacciones en el que la primera instrucción de cada transacción marca su inicio y donde los t_i son instantes de tiempo ($t_i < t_j$ si $i < j$):



Suponiendo que se aplica el protocolo OMT:

- a) ¿Cuál es el plan en serie al que debería ser equivalente el plan concurrente anterior? **T1→T2→T3**
- b) Si antes del inicio del plan, tanto MT_E(X) como MT_L(X) (marcas de tiempo de escritura y de lectura de X) valen t_{-20} donde $t_{-20} < t_0$, comprueba si el protocolo **OMT** admite el plan dando una traza de su ejecución. Usa la tabla para contestar esta pregunta. **Si en algún instante el plan ya no es posible di por qué y explica qué condición ha permitido detectar el problema.**

				X	
t	T1	T2	T3	ML	ME
				t_{-20}	t_{-20}
t_0	$w_1(X)$			t_{-20}	t_0
t_1		$r_2(X)$		t_1	t_0
t_2		$w_2(X)$		t_1	t_1
t_3			$r_3(X)$	t_3	t_1
t_4		$r_2(X)$		t_3	t_1
t_5			$w_3(X)$	t_3	t_3
t_6	$r_1(X)$				
t_7					
t_8					
t_9					

La operación $r_1(X)$ de T_1 se aborta ya que $MT(T1) < MT_E(X)$ ($t_0 < t_3$). T_1 debe revertirse e introducirse de nuevo en el sistema.

En t_6 el protocolo ha detectado que T_1 está intentando leer un dato modificado por una transacción que en el plan en serie equivalente cronológico es más joven que ella.

3. (0,25 puntos)

Supón un entorno concurrente controlado por el protocolo Multiversión que hay tres versiones del elemento x (sean x_0 , x_1 y x_2) cuyas marcas de tiempo son las que se indican más abajo (donde $t_i < t_j$ si $i < j$), en un momento determinado la transacción T con marca de tiempo $MT(T) = t_4$ ejecuta la instrucción $w(x)$ que es rechazada.

	MT_L	MT_E
x_0	t_5	t_2
x_1	t_5	t_5
x_2	t_8	t_8

- a) ¿Qué condición o condiciones se han evaluado para detectar que la operación deba rechazarse?
- b) ¿Qué se controla con esa condición?
- c) ¿Qué marca de tiempo debería tener T para que la operación fuera aceptada y cuáles serían las marcas de tiempo de la versión creada? (puede haber varias respuestas correctas)
- a) De las tres versiones de x que hay se busca la que tenga mayor marca de tiempo de escritura que cumpla que: $MT_E(x_i) \leq MT(T)$. En este caso esa versión es x_0 . Dado que $MT_L(x_0) > MT(T)$, la operación es rechazada ya que x_0 ha sido leída por una transacción más joven que T .
- b) Con esta condición se controla que la versión “anterior” al elemento de datos x no ha sido leída por ninguna transacción T^* posterior (cronológicamente) a T .
- c) Hay varias posibilidades:
- $MT(T) = t_6$ La versión creada x_4 tendría $MT_L(x_4) = t_6$ y $MT_E(x_4) = t_6$
 - $MT(T) = t_9$ La versión creada x_4 tendría $MT_L(x_4) = t_9$ y $MT_E(x_4) = t_9$

Cada una de las 20 cuestiones siguientes vale 0,1 puntos. Cada cuestión mal contestada resta 1/3 de una respuesta correcta. Nota Test = $(0,1 \times \text{Respuestas_correctas}) - ((0,1/3) \times \text{Respuestas_incorrectas})$. **Puedes usar el espacio en blanco al final de la página como espacio en sucio.**

- 1) Sea F un fichero con 10^6 registros que tiene tres atributos A , B y C tal que A es clave y B y C no lo son, ¿cuál de las siguientes afirmaciones es FALSA?
 - a) No es lo mismo definir un índice sobre (B, C) que definirlo sobre (C, B) .
 - b) F podría tener a la vez un índice primario sobre A y un índice de agrupación sobre B .
 - c) Un índice disperso sobre F tendría menos de 10^6 entradas.
 - d) F podría tener a la vez un índice secundario sobre B y un índice secundario sobre C .
- 2) ¿Cuál de las siguientes afirmaciones es FALSA?
 - a) Un índice es denso si tiene la misma cantidad de entradas que registros tiene el fichero de datos que indexa.
 - b) En un índice multinivel, el primer nivel puede ser un índice primario, secundario o de agrupación.
 - c) Un índice puede construirse como un fichero ordenado o como un árbol de búsqueda.
 - d) En un índice multinivel de altura 4, la búsqueda de un valor en el índice puede suponer el acceso a menos de cuatro bloques del índice.
- 3) ¿Qué significa que la ocupación de los nodos en un índice en árbol es uniforme?
 - a) Que todos los nodos hojas están al mismo nivel.
 - b) Que la ocupación de todos los nodos es al menos del 50%.
 - c) Que el coste de localizar la dirección de un registro dado un valor del campo de indexación es siempre el mismo.
 - d) Que el árbol es independiente del orden de llegada de los valores.
- 4) Respecto a los índices en árbol B^+ , ¿cuál de las siguientes afirmaciones es FALSA?
 - a) Todos los valores del campo de indexación del fichero aparecen en las hojas del índice.
 - b) Los nodos hoja están doblemente enlazados entre sí.
 - c) Los punteros a nodos del árbol aparecen en los nodos internos y en los nodos hoja.
 - d) Si el orden del índice es p , en un nodo hoja hay p valores del campo de indexación y $p-1$ punteros a registro del fichero indexado.
- 5) Sea F un fichero con 200.000 registros ordenado por A y sea 10 el factor de bloque del fichero. Supón que el fichero tiene un índice sobre el campo A , ¿cuántas entradas tendrá el índice según A sea clave o no?
 - a) 200.000 si A es clave y 2.000 si A no es clave y cada valor de A se repite 10 veces.
 - b) 200.000 si A es clave y 20.000 si A no es clave y cada valor de A se repite 10 veces.
 - c) 20.000 si A es clave y 20.000 si A no es clave y cada valor de A se repite 10 veces.
 - d) 20.000 si A es clave y 2.000 si A no es clave y cada valor de A se repite 10 veces.

6) ¿Qué afirmación sobre el fichero disperso es FALSA?

- a) Asegura la búsqueda más económica, con condición de rango sobre el campo de dispersión.
- b) Exige conocer de antemano el tamaño estimado que va a tener el fichero.
- c) No es adecuado para hacer búsquedas ordenadas por el campo de dispersión.
- d) Si hay muchas colisiones la eficiencia de las búsquedas por igualdad en el campo de dispersión disminuye.

7) Dada una operación de búsqueda con condición de selección $A \geq 9$ en un fichero desordenado, ¿qué índice sería el más adecuado?

- a) Índice B+ sobre A.
- b) Índice B sobre A.
- c) Los índices sólo son útiles con condiciones de igualdad.
- d) Índice primario sobre A.

8) Sea F un fichero con dos campos A (clave) y B y se quiere definir un índice sobre A ¿cuál de las siguientes afirmaciones es CIERTA?

- a) En lugar de definir el fichero F y el índice, se puede organizar todo en una única estructura de índice donde para cada valor de A en lugar de un puntero a registro se almacene el valor de B asociado a ese A.
- b) Se puede definir el fichero F y el índice donde para cada valor de A en lugar de un puntero a registro se almacene el valor de B asociado a ese A.
- c) En lugar de definir el fichero F y el índice, se puede organizar todo en una única estructura de índice donde para cada valor de A además de un puntero a registro se almacene el valor de B asociado a ese A.
- d) En este caso no se puede organizar todo en una única estructura por ser A clave.

9) ¿Cuál de las siguientes afirmaciones es CIERTA?

- a) En el peor de los casos, la inserción o eliminación de registros tiene el mismo coste en las tres organizaciones de ficheros.
- b) En el mejor de los casos, la inserción o eliminación de registros tiene el mismo coste en las tres organizaciones de ficheros.
- c) En un fichero disperso el coste de modificar un campo es siempre eficiente sea cual sea el campo.
- d) Los cubos de desbordamiento en un fichero disperso agilizan las búsquedas ordenadas en el fichero.

10) ¿Cuál de las siguientes afirmaciones es CIERTA?

- a) Un registro es la unidad de transferencia de datos entre memoria principal y memoria secundaria.
- b) Un registro es la unidad de direccionamiento de datos.
- c) Desde un punto de vista lógico, un fichero es una secuencia de registros.
- d) Un bloque se puede almacenar en dos discos distintos.

- 11) ¿Cuál será la mejor implementación de una tabla cuya clave primaria está formada por todas sus columnas y cuya forma de manipulación no se conoce?
- Como un fichero ordenado, para poder realizar búsquedas rápidas y poder comprobar la restricción de unicidad eficientemente.
 - Como un fichero desordenado para permitir una manipulación flexible y con un índice sobre las columnas de la clave primaria para poderla comprobar eficientemente.
 - Como un fichero disperso, con función de dispersión sobre la clave primaria.
 - Organizada como un índice.**
- 12) Si $\langle P_1, K_1, P_2, K_2, \dots, P_{q-1}, K_{q-1}, P_q \rangle$ representa un nodo de un índice en árbol B, ¿qué condición se cumple para todos los valores X del subárbol apuntado por el puntero P_2 ?
- $K_1 < X < K_2$**
 - $K_1 < X \leq K_2$
 - $K_1 \leq X < K_2$
 - $K_1 \leq X \leq K_2$
- 13) ¿Cuál de las siguientes reglas NO es una regla heurística para la optimización de consultas?
- Mover las selecciones hacia abajo en las ramas del árbol.
 - Mover las proyecciones hacia arriba en las ramas del árbol.**
 - Combinar un producto cartesiano con una selección para obtener una concatenación.
 - Descomponer las selecciones con condiciones conjuntivas en una secuencia de selecciones simples.
- 14) Sea la siguiente expresión de Álgebra Relacional: $\pi_{A,B}(\pi_{A,B,C}(\sigma_{C<7 \wedge B=5 \wedge A=1}(R)))$, ¿a cuál de las siguientes expresiones es equivalente después de aplicar tantas veces como sea posible la regla de transformación1 (*Cascada de σ*)?:
- $\pi_{A,B}(\sigma_{C<7}(\sigma_{(B=5 \wedge A=1)}(R)))$
 - $\pi_{A,B,C}(\sigma_{C<7}(\sigma_{(B=5 \wedge A=1)}(R)))$
 - $\pi_{A,B}(\sigma_{C<7}(\sigma_{B=5}(\sigma_{A=1}(R))))$**
 - $\pi_{A,B,C}(\sigma_{C<7}(\sigma_{B=5}(\sigma_{A=1}(R))))$
- 15) Sobre los métodos de búsqueda para resolver una selección con condición simple sobre una tabla, ¿cuál de las siguientes afirmaciones es FALSA?:
- La búsqueda lineal se puede aplicar siempre sea cual sea la condición de la selección.
 - La búsqueda binaria se puede aplicar siempre sea cual sea la condición de la selección.**
 - La búsqueda con direccionamiento calculado sólo se puede aplicar a condiciones de igualdad.
 - La búsqueda con índice sólo se puede aplicar a condiciones de igualdad y de rango.

- 16) ¿Sea R una tabla con 1.000 tuplas y sea la operación del álgebra relacional: $\sigma_{A=7} R$, ¿cuál de las siguientes afirmaciones es FALSA sobre la selectividad de la condición $A=7$, $sl_{A=7}$?
- a) Si A es la clave primaria de R , entonces $sl_{A=7} = 1/1.000$
 - b) Sirve para que el sistema pueda estimar cuántas filas va a tener una selección.
 - c) Si en R , en el atributo A aparecen 20 valores distribuidos uniformemente, entonces $sl_{A=7} = 1/20$
 - d) El SGBD en todo momento conoce con exactitud el valor de $sl_{A=7}$
- 17) ¿Qué métodos hay para computar un operador en un árbol de consulta?
- a) Materializada e indexada
 - b) Indexada y en flujo
 - c) Materializada y en flujo
 - d) Materializada, indexada y en flujo.
- 18) Sobre los árboles de consulta, ¿cuál de las siguientes afirmaciones es FALSA?
- a) Se evalúan de arriba abajo.
 - b) La ejecución de un nodo implica sustituir el nodo por el resultado de la operación del nodo.
 - c) Un nodo con un operador binario tendrá siempre dos hijos.
 - d) Los nodos hoja son tablas del esquema.
- 19) En el diseño físico de una base de datos, ¿cuál de las siguientes acciones no es adecuada para mejorar el rendimiento de la base de datos?:
- a) Desnormalización.
 - b) Partición de tablas (horizontal o verticalmente).
 - c) Fusión de tablas.
 - d) Eliminación de claves primarias.
- 20) Respecto a la implementación estándar de una base de datos relacional, ¿cuál de las siguientes afirmaciones es FALSA?
- a) Es útil definir índices sobre las claves primarias y sobre conjuntos con restricción de unicidad.
 - b) Es útil definir índices sobre las claves ajenas.
 - c) Es útil definir índices sobre atributos que pueden ser nulos y que se usan en condiciones de desigualdad.
 - d) En general, una organización desordenada es adecuada para cualquier tabla si no se conoce ni el tamaño que va a tener ni las consultas que se van a realizar sobre ella.