Tema.- Gestión de transacciones

Fundamentos de Sistemas de Bases de Datos. Ramez A. Elmasri, Shamkant B. Navathe [cap. 19]

Índice

Transacción Atómica

- Unidad lógica de procesamiento de BD
- Propiedades deseables (Atomicidad, Consistencia, Aislamiento, Durabilidad)
- Planes (historias) de transacciones, Recuperabilidad

Control Concurrencia

- Varias transacciones concurrentes se intercalan
- Seriabilidad: secuencias de ejecución correctas

Recuperación ante fallos por:

- error de transacción (overflow, división por 0, interrupción del usuario, ...)
- excepciones de transacción (NO ∃ datos, saldo insuficiente, ...)
- control concurrencia (viola la seriabilidad, bloqueo mortal, ...)
- caída del sistema
- I fallo del disco
- catástrofes físicas

Procesamiento de Transacciones

- Transacción Atómica
 - Unidad lógica de procesamiento de BD
 - Ejecución de un programa que LEE o ESCRIBE en la BD Operaciones de Acceso:

```
leer_elemento (X) \Rightarrow X (disco) \rightarrow memoria \rightarrow X (variable)
```

```
escribir_elemento (X) \Rightarrow X (disco)
memoria \rightarrow X (disc\mathfrak{F})
X (variable)
```

EJ: RESERVAS DE VUELOS:

T1: transfiere N reservas del vuelo X al vuelo Y

leer_elemento (X)

X:= X - N

escribir_elemento(X)

leer_elemento(Y)

Y:= Y + N

escribir_elemento(Y)

Conjunto lectura = {X,Y}

Conjunto escritura = $\{X,Y\}$

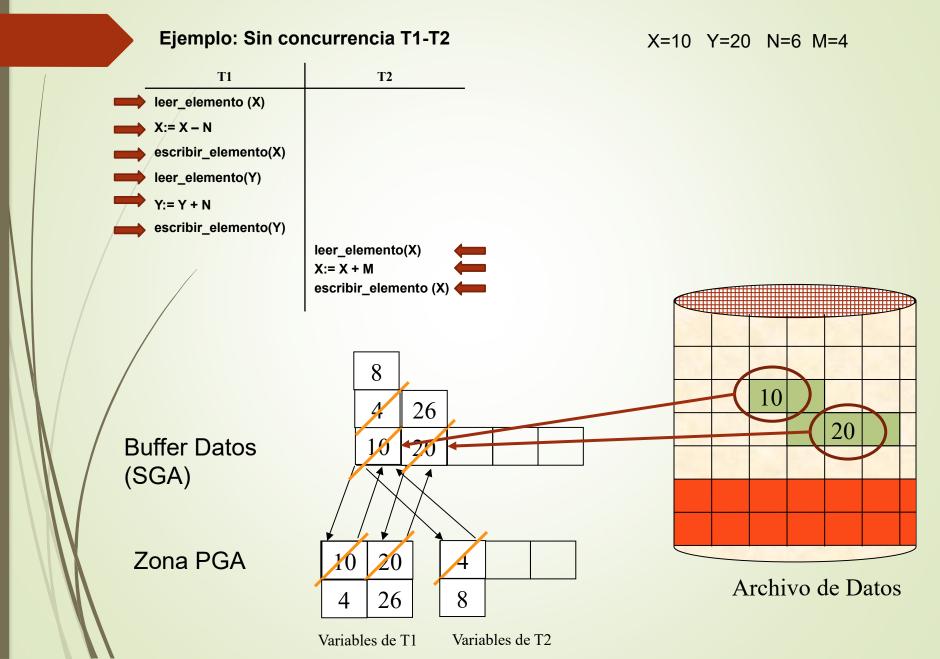
T2: añade M reservas al vuelo X

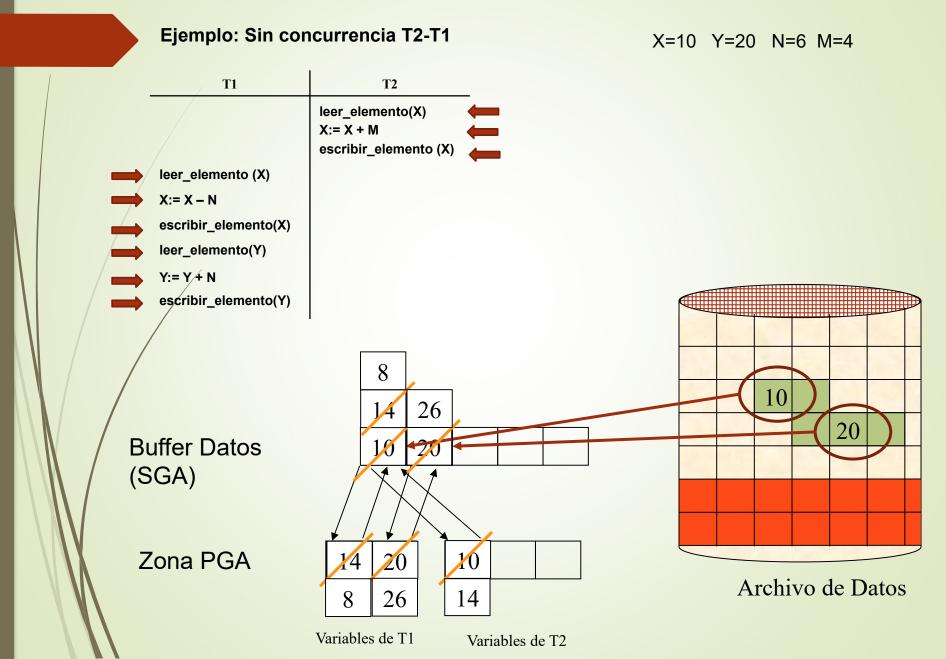
leer_elemento (X)

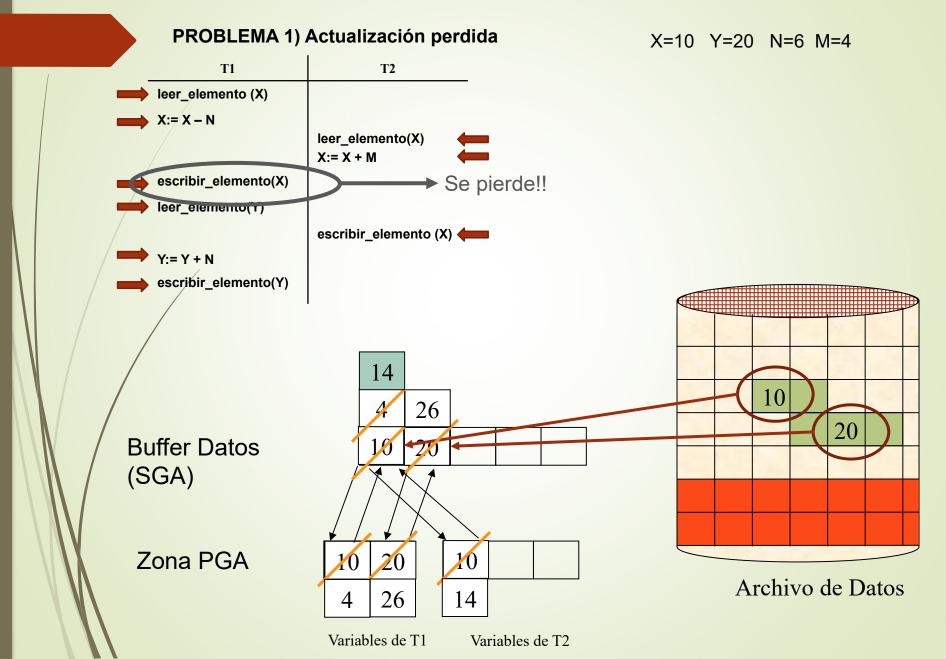
X:= X + M

escribir_elemento(X)

Conjunto lectura = {X}
Conjunto escritura = {X}





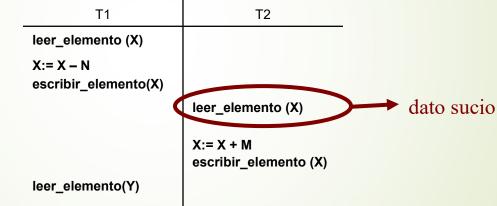


EJ: RESERVAS DE VUELOS:

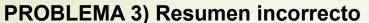
T1: transfiere N reservas del vuelo X al vuelo Y

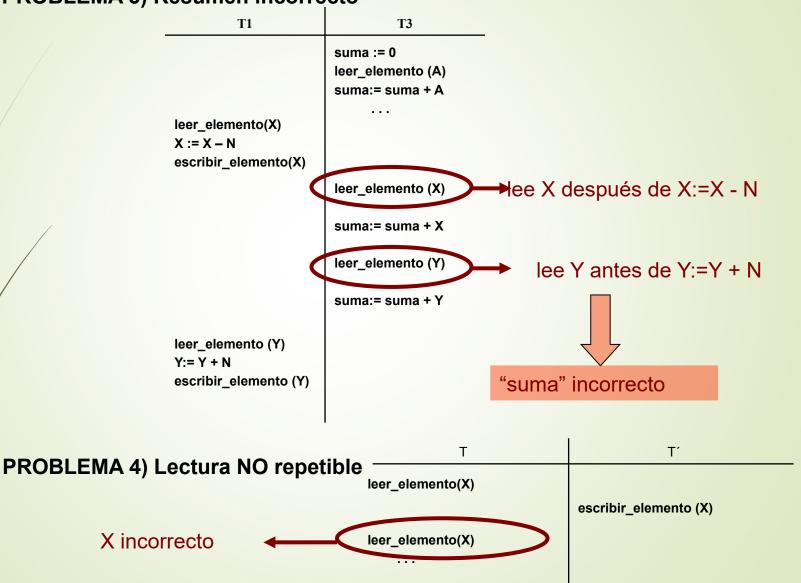
T2: añade M reservas al vuelo X

PROBLEMA 2) Actualización temporal (lectura sucia)

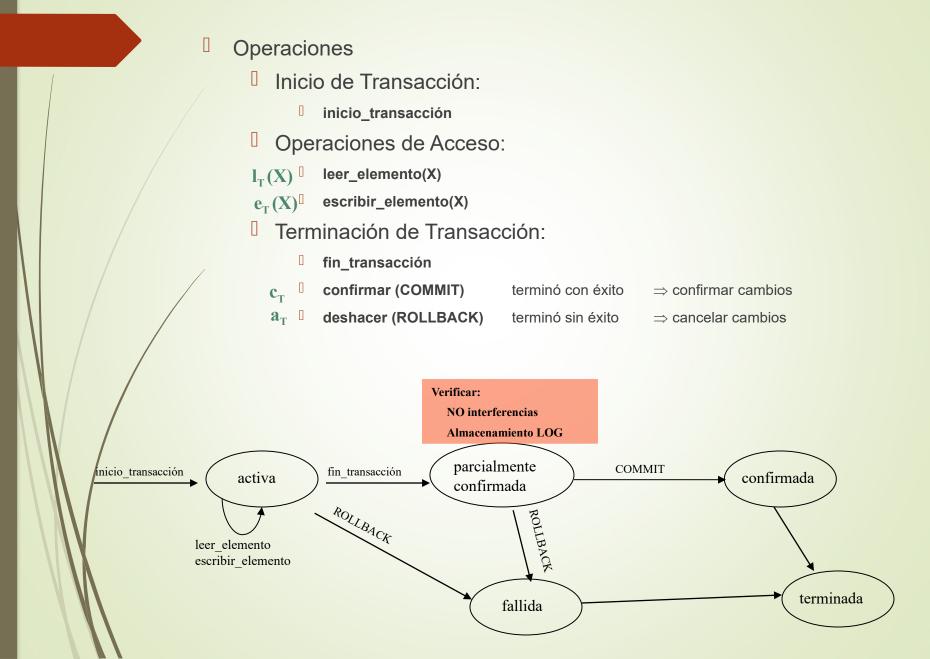


T1 falla ⇒ X debe volver a su antiguo valor PERO T2 ya ha leído X





Estado de una transacción



Archivo de Log

= {registros Log} que se mantienen temporalmente en memoria Tipos de Registros Log:

```
      [inicio_transacción, T]
      inicio ejecución de T

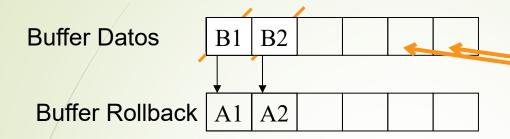
      [escribir_elemento, T, X, valor_ant, valor_nuevo]
      cambio de X(disco)

      [leer_elemento, T, X]
      lectura del elemento X (disco)

      [confirmar, T]
      ∀ escrituras se han ejecutado con éxito ∀ reg. Log de T→ A. Log (forzar la escritura)

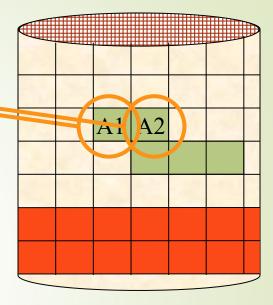
      [abortar, T]
      falló o se abortó la transacción T ∀ reg. Log de T→ A. Log (forzar la escritura)
```

ZONA DE MEMORIA (SGA)

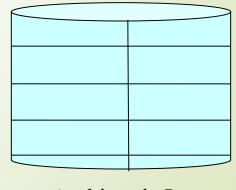


Tr.	Fecha	Estado	Dirección	Viejo	Nuevo
T1		En curso	tabla1.D1	A1	B1
T2		En curso	tabla2.D2	A2	B2

- 1) T1 cambia el dato D1=A1 al valor B1
- 2) T2 cambia el dato D2=A2 al valor B2



Archivo de Datos



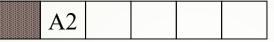
Archivo de Log

ZONA DE MEMORIA (SGA)

Buffer Datos

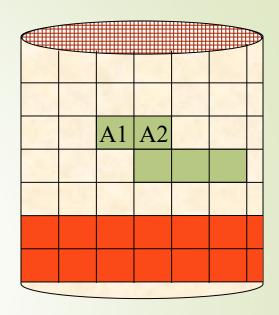
B1 B2				
-------	--	--	--	--

Buffer Rollback

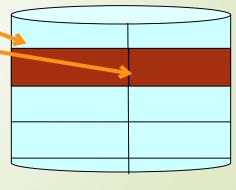


Tr.	Fecha	Estado	Dirección	Viejo	Nuevo
T1 /		En curso	tabla1.D1	A1	B1
T2		En curso	tabla2.D2	A2	B2
/ T1		Confirmar			

- 1) T1 cambia el dato D1=A1 al valor B1
- 2) T2 cambia el dato D2=A2 al valor B2
- 3) T1 confirmado
 - => escribir Reg. Log
 - => mover Reg. Log de T1 a Archivo Log
 - => liberar Buffer Rollback



Archivo de Datos



Archivo de Log

ZONA DE MEMORIA (SGA)

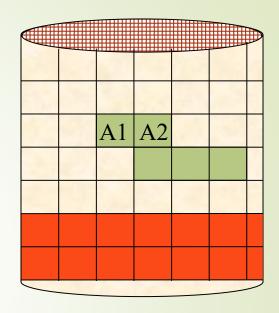
Buffer Datos

B1 A2

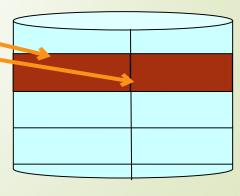
Buffer Rollback A1 A2

Tr.	Fecha	Estado	Dirección	Viejo	Nuevo
T1 /		En curso	tabla1.D1	A1	B1
T2/		En curso	tabla2.D2	A2	B2
/Γ2		Abortar			

- 1) T1 cambia el dato D1=A1 al valor B1
- 2) T2 cambia el dato D2=A2 al valor B2
- 3) T2 anulado
 - => escribir Reg. Log
 - => mover Reg. Log de T2 a Archivo Log
 - => mover Buffer Rollback a Buffer Datos



Archivo de Datos

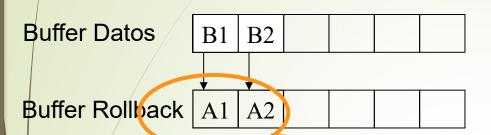


Archivo de Log

Archivo de Log

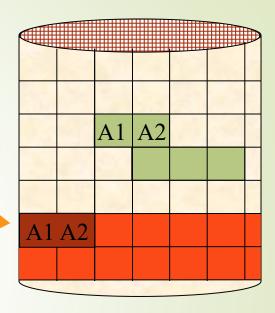
```
Tipos de Registros Log:
[inicio_transacción, T] inicio ejecución de T
[escribir elemento, T, X, valor ant, valor nuevo] cambio de X(disco)
[leer_elemento, T, X] lectura del elemento X (disco)
[confirmar, T]
                            ∀ escrituras se han ejecutado con éxito
                      \forall reg. Log de T\rightarrow A. Log (forzar la escritura)
[abortar, T]
                            se abortó la transacción T
                      \forall reg. Log de T\rightarrow A. Log (forzar la escritura)
[checkpoint] (punto de control) \forall datos (memoria) de T confirmadas \rightarrow disco
                      \forall reg. Log \rightarrow A. Log (forzar la escritura)
   se realiza periódicamente (m minutos, N transacciones confirmadas)
   almacena una lista con T activas en el instante checkpoint
   SI FALLO ⇒ T terminadas antes del último checkpoint NO se repiten
           T terminadas después se repiten (REDO) ⇒ rastrear A. Log hacia adelante
                                  cambiando X al valor nuevo
           T NO terminadas se anulan (UNDO) ⇒ rastrear A. Log hacia atrás
                                  restableciendo X al valor ant
```

ZONA DE MEMORIA (SGA)

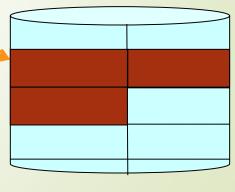


	Tr.	Fecha	Estado	Dirección	Viejo	Nuevo
	T1		En curso	tabla1.D1	A1	B1
	T2		En curso	tabla2.D2	A2	B2
I	CHECKPOINT					

- 1) T1 cambia el dato D1=A1 al valor B1
- 2) T2 cambia el dato D2=A2 al valor B2
- 3) CHECKPOINT
 - => escribir Reg. Log
 - => mover Buffer Rollback a disco
 - => mover Buffer de Datos de trans.confirmadas a disco (no hay)
 - => mover Reg. Log a Archivo Log



Archivo de Datos



Archivo de Log

Propiedades deseables de las transacciones

Atomicidad (Atomicity)

- ∀ operaciones o ninguna
- es responsabilidad del MECANISMO DE RECUPERACIÓN

Consistencia (Consistency)

- pasar de un estado consistente (cumple ∀ restricciones) a otro
- es responsabilidad del programador que las transacciones sean independientes

Aislamiento (Isolation)

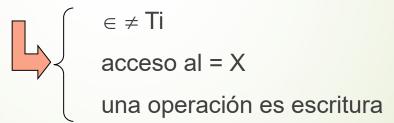
la ejecución de una transacción no debe interferir en otras transacciones concurrentes lo impone el MECANISMO DE CONTROL DE CONCURRENCIA

Permanencia (Durability)

- las transacciones confirmadas NO se pierden
- es responsabilidad del MECANISMO DE RECUPERACIÓN

Planes de transacciones

- Orden de ejecución de las operaciones ⊂ {T1, T2, ..., Tn}
 - □ ⊂ ∀ operaciones (incluso [confirmar, T] [abortar, T])
 - = orden de las operaciones que en Ti
 - si ∃ operaciones en <u>conflicto</u> ⇒ secuencial



Planes según su recuperabilidad

1) RECUPERABLE ⇒ T confirmada nunca se tiene que deshacer (rollback)

Ti no finaliza mientras no hayan finalizado las últimas Tj que escriben antes elementos leídos por Ti

ej:
$$I_1(X)$$
 $e_1(X)$ $I_2(X)$ $I_1(Y)$ $e_2(X)$ $e_1(Y)$
ej: $I_1(X)$ $e_1(X)$ $I_2(X)$ $I_1(Y)$ $e_2(X)$ $e_1(Y)$

GARANTIZA NO anulación de Ti confirmadas PERO... SI anulación de Ti NO confirmadas

ej:
$$I_1(X)$$
 $e_1(X)$ $I_2(X)$ $I_1(Y)$ $e_2(X)$ $e_1(Y)$ a_1 a_2 anulación en cascada

PERO...

2) ÉVITA ANULACIÓN EN CASCADA

Ti no lee X mientras no hayan finalizado las Tj que escriben X antes ej: l₂(X) (y todas las operaciones posteriores en T2) se pospone hasta a₄ ó c₄

3) PLAN ESTRICTO

Ti no lee ni escribe X mientras no hayan finalizado las Tj que escriben X antes

ej:
$$e_1(X)$$
 $e_2(X)$ a_1

Sol.- $e_2(X)$ se pospone hasta a_1 ó c_1

→ No Plan Estricto restauración de la imagen anterior (valor de X antes de Ti abortada)

Ejemplos planes según su recuperabilidad

Para cada uno de los planes siguientes, indicar si son recuperables (RE), evitan anulación en cascada (AC), estrictos (ES), del siguiente modo:

- En caso afirmativo escribir S en el recuadro correspondiente
- Si no se puede decir si un plan pertenece a cierta clase según las acciones indicadas, explicarlo brevemente (para ello, en el recuadro correspondiente poner un número y explicar debajo de la tabla).

Las acciones aparecen en el orden consecutivo dentro de cada plan. En caso de que no aparezcan las confirmaciones de las transacciones, asumir que los *commit l abort* se encuentran después de todas las acciones indicadas.

	PLAN	RE	AC	ES
1	I1(X) I2(X) e1(X) e2(X)			
2	e1(X) I2(Y) I1(Y) I2(X)			
3/	I1(X) I2(Y) e3(X) I2(X) I1(Y)			
4	I1(X) I1(Y) e1(X) I2(Y) e3(Y) e1(X) I2(Y)			
5	I1(X) e2(X) e1(X) a2 c1			
6	I1(X) e2(X) e1(X) c2 c1			
7	I2(X) e3(X) c3 e1(Y) c1 I2(Y) e2(Z) c2			
8	I1(X) e2(X) c2 e1(X) c1 I3(X) c3			

Ejemplos planes según su recuperabilidad

RECUPERABLE: Ti **NO FINALIZA** mientras no hayan <u>finalizado</u> las últimas Tj que escriben antes elementos leídos por Ti EVITA ANULACIÓN EN CASCADA: Ti **NO LEE** X mientras no hayan <u>finalizado</u> las Tj que escriben X antes

PLAN ESTRICTO: Ti NO LEE NI ESCRIBE X mientras no hayan finalizado las Tj que escriben X antes

	PLAN	RE	AC	ES
1	I1(X) I2(X) e1(X) e2(X)			
2	e1(X) I2(Y) I1(Y) I2(X)			
3	I1(X) I2(Y) e3(X) I2(X) I1(Y)			
4	I1(X) I1(Y) e1(X) I2(Y) e3(Y) e1(X) I2(Y)			
5	I1(X) e2(X) e1(X) a2 c1			
6	11(X) e2(X) e1(X) c2 c1			
7	I2(X) e3(X) c3 e1(Y) c1 I2(Y) e2(Z) c2			
8	I1(X) e2(X) c2 e1(X) c1 I3(X) c3			

- (a) se cumple si T1 finaliza antes de T2
- (b) se cumple si T3 finaliza antes de T2
- (c) se cumple si T3 finaliza antes de T2

Seriabilidad de los planes

Determina qué planes son "correctos"

1) PLAN EN SERIE:

- operaciones de T se ejecutan consecutivamente (sin intercalación)
- Si Ti's *independientes* (consistencia) ⇒ ∀ plan serie es CORRECTO
- Problema: limita la concurrencia

2) PLAN NO EN SERIE:

- Intercala operaciones entre Ti's
- SERIALIZABLE ⇒ EQUIVALENTE a un PLAN EN SERIE ⇒ correcto



por resultados ⇒ produce el mismo estado en la BD para cada dato se aplican = operaciones y en = orden

2.1) por conflictos 2.2 por vistas

Ejemplo:

P1
leer_elemento (X)
X:= X + 10
escribir_elemento(X)

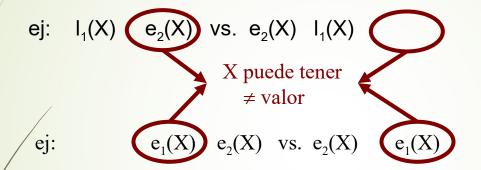
P2
leer_elemento (X)
X:= X * 1.1
escribir_elemento(X)

equivalentes por resultado si X=100

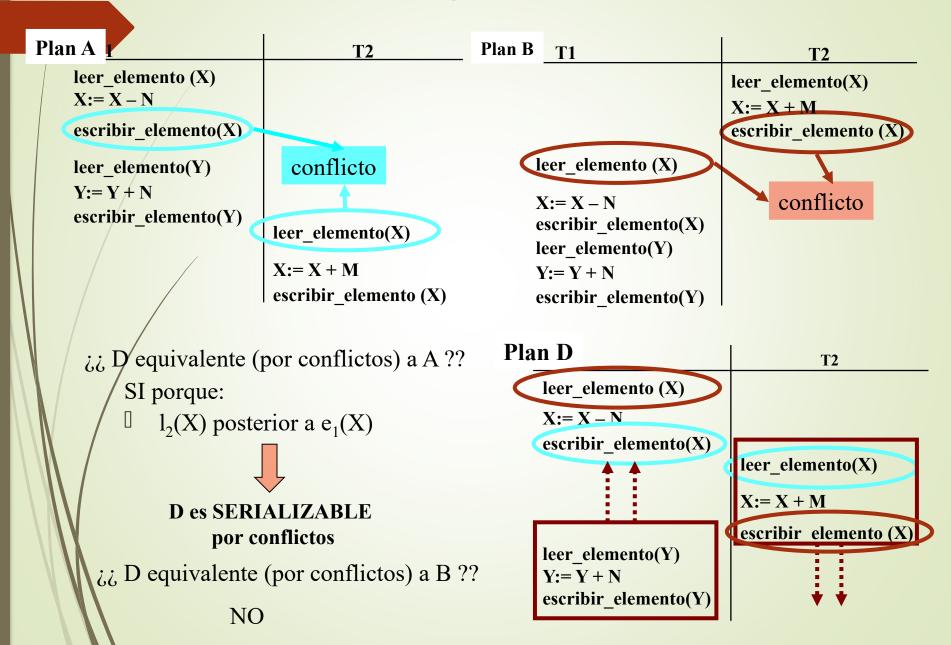
Seriabilidad de los planes

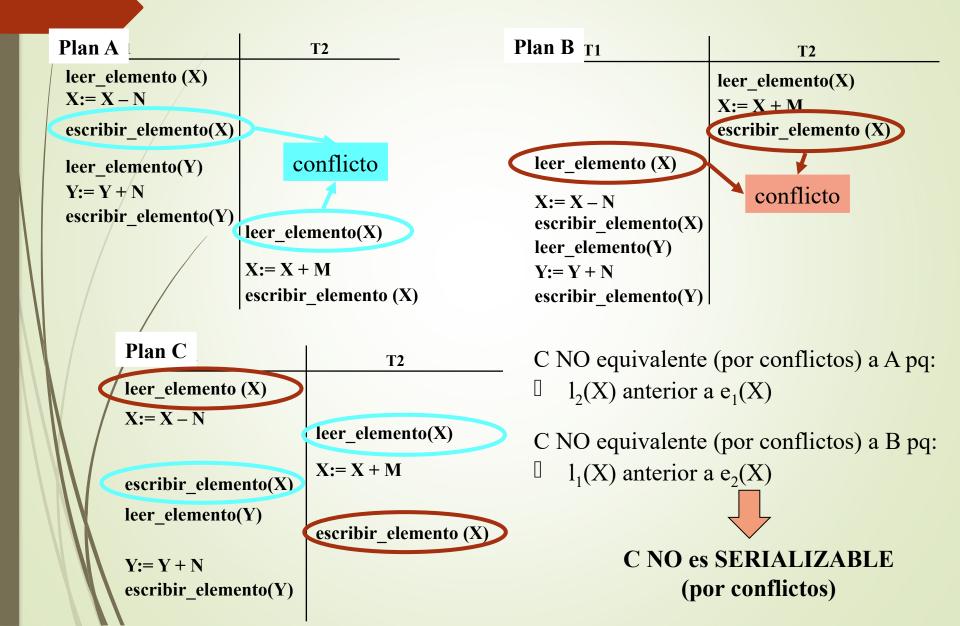
2.1) 2 planes son EQUIVALENTES POR CONFLICTOS:

El orden de las operaciones en conflicto es = en ambos planes



P NO en serie es SERIALIZABLE POR CONFLICTOS ⇒
es EQUIVALENTE POR CONFLICTOS a P'en serie



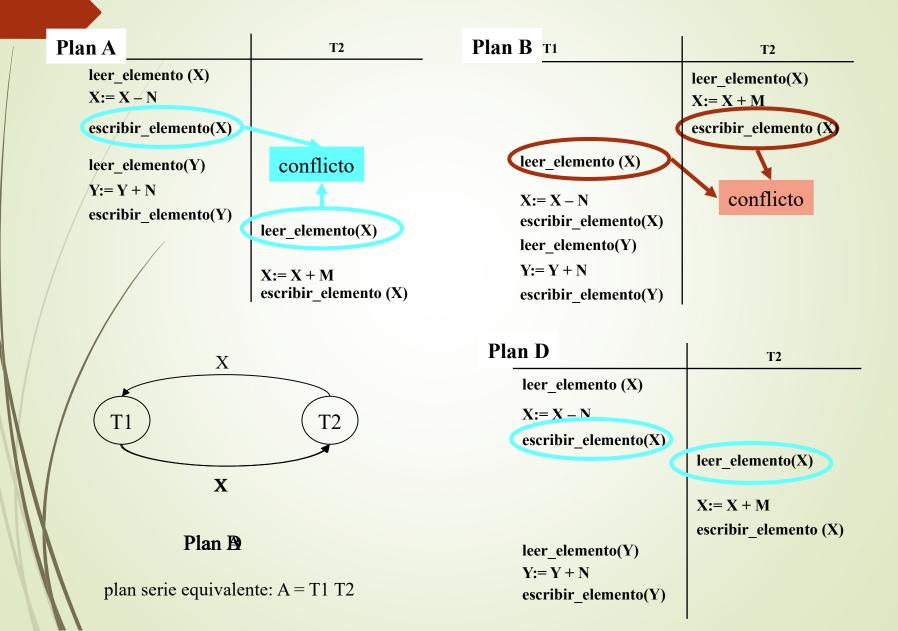


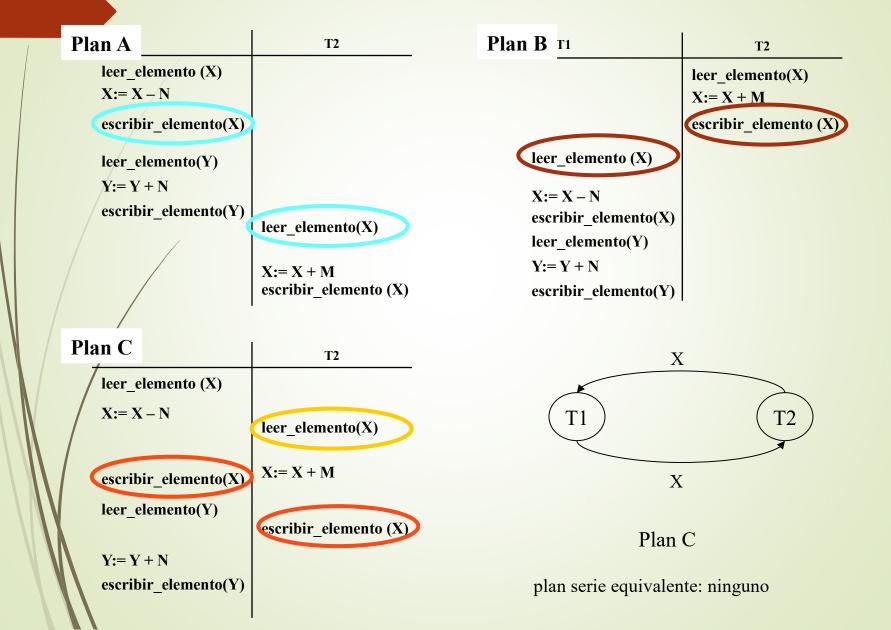
Prueba de Seriabilidad por Conflictos

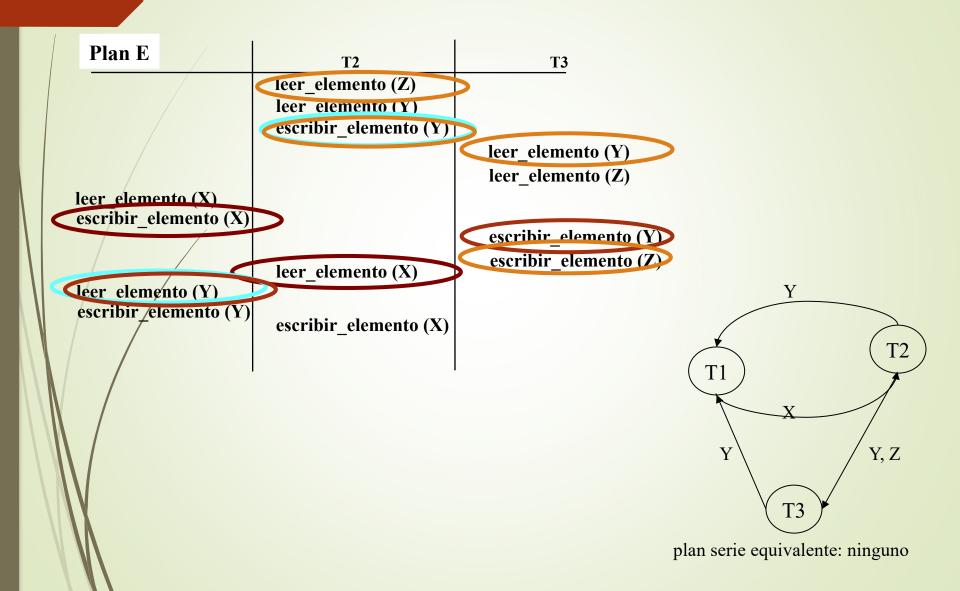
- Construir un GRAFO DE PRECEDENCIA G= (N, A) siendo:
 - $N = \{T1, T2, ..., Tn\}$
 - A = $\{a1, a2, ..., am\}$ donde ap = $(Ti \rightarrow Tj)$
 - ⇒ 1 oper. de Ti aparece antes que una operación en conflicto de Tj
 - ⇒ en un plan en serie Ti debe aparecer antes de Tj

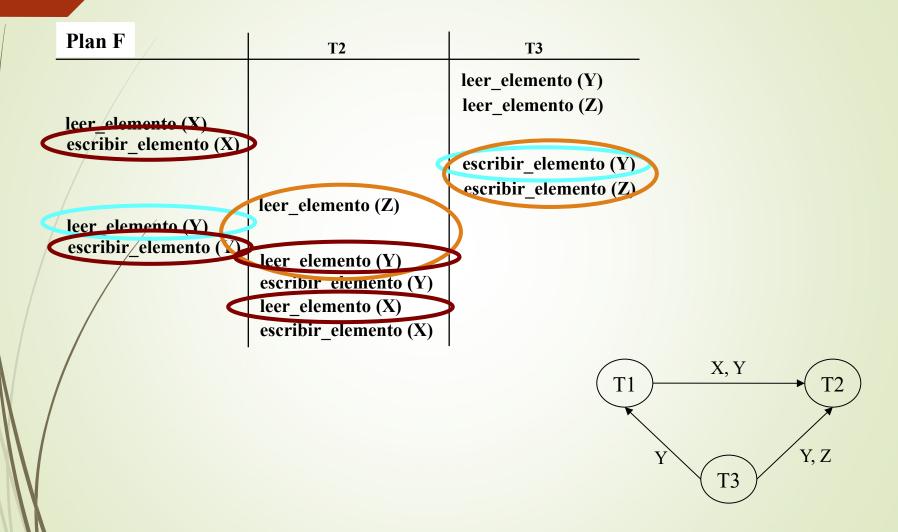
ALGORITMO:

- 1. crear un nodo ∀ Ti de P
- 2. ... $e_i(X)$... $l_j(X)$... $\Rightarrow Ti \rightarrow Tj$
- 3. ... $l_i(X)$... $e_i(X)$... $\Rightarrow Ti \rightarrow Tj$
- 4. ... $e_i(X)$... $e_i(X)$... $\Rightarrow Ti \rightarrow Tj$
- 5. P es SERIALIZABLE (por conflictos) ⇔ el grafo NO tiene ciclos
- 6. el plan en serie equivalente (por conflictos) se obtiene siguiendo las aristas (ordenación topológica)









plan serie equivalente: $T3 \rightarrow T1 \rightarrow T2$

Seriabilidad de los planes

2.2) P y P'son EQUIVALENTES POR VISTAS ⇒

1.
$$\forall$$
 Ti \in P \Rightarrow Ti \in P'

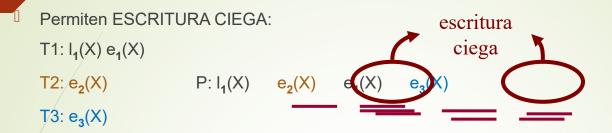
2.
$$\begin{cases} I_{\mathbf{i}}(X) \\ e_{\mathbf{j}}(X) I_{\mathbf{i}}(X) \end{cases} = \begin{bmatrix} I_{\mathbf{i}}(X) \\ e_{\mathbf{j}}(X) I_{\mathbf{i}}(X) \end{bmatrix} = \begin{bmatrix} I_{\mathbf{i}}(X) \\ e_{\mathbf{j}}(X) I_{\mathbf{j}}(X) \end{bmatrix} = \begin{bmatrix} I_{\mathbf{i}}(X) \\$$

(∀ lectura lee el resultado de la misma escritura en ambos planes, es decir, P y P'perciben la MISMA VISTA)

3. $e_k(Y)$ última escritura de Y en P \Rightarrow $e_k(Y)$ última escritura de Y en P' (MISMO ESTADO FINAL de la BD)

P NO en serie es SERIALIZABLE POR VISTAS ⇒
es EQUIVALENTE POR VISTAS a P'en serie

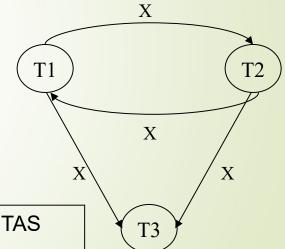
Planes Serializables por Vistas



P es SERIALIZABLE POR VISTAS (equivalente por vistas a P' = {T1, T2, T3})

```
T1T2T3: I_1(X) e_1(X) e_2(X) e_3(X)
porque I_1(X) es limpia en ambos planes y
e_3(X) es la última escritura en ambos planes
```

perø NO ES SERIALIZABLE POR CONFLICTOS



∀ P serializable por CONFLICTOS es serializable por VISTAS

```
\forall P serializable por VISTAS es serializable por CONFLICTOS \Leftrightarrow \forall e<sub>i</sub>(X) \Rightarrow \exists I<sub>i</sub>(X), y depende SOLO de I<sub>i</sub>(X) (ESCRITURA RESTRINGIDA)
```

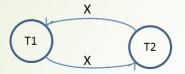
Indicar si los planes siguientes son serializables por conflictos (SC) y/o serializables por vistas (SV) Si no se puede decir si un plan pertenece a cierta clase según las acciones indicadas, explicarlo brevemente.

Las acciones aparecen en el orden consecutivo dentro de cada plan. En caso de que no aparezcan las confirmaciones de las transacciones, asumir que los *commit | abort* deberán encontrarse después de todas las acciones indicadas.

	PLAN	SC	SV
1	I1(X) I2(X) e1(X) e2(X)		
2	e1(X) I2(Y) I1(Y) I2(X)		
3	I1(X) I2(Y) e3(X) I2(X) I1(Y)		
4	I1(X) I1(Y) e1(X) I2(Y) e3(Y) e1(X) I2(Y)		
5	I1(X) e2(X) e1(X) a2 c1		
6	I1(X) e2(X) e1(X) c2 c1		
7	I2(X) e3(X) c3 e1(Y) c1 I2(Y) e2(Z) c2		
8	I1(X) e2(X) c2 e1(X) c1 I3(X) c3		

P: I1(X) I2(X) e1(X) e2(X)

¿SERIALIZABLE POR CONFLICTOS?



Conclusión:

Existe un ciclo, por lo que no es serializable por conflictos

¿SERIALIZABLE POR VISTAS?

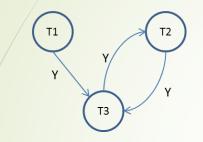
T1T2: I1(X) e1(X) l2(X) e2(X). No es equivalente, dado que en el plan en serie la lectura I2(X) es sucia, mientras que en el plan a comparar es limpia.

T2T1: I2(X) e2(X) I1(X) e1(X). No es equivalente, dado que en el plan en serie la lectura I1(X) es sucia, mientras que en el plan a comparar es limpia.

Conclusión: **NO es serializable por vistas**, dado que no existe ningún plan en serie equivalente a P por vistas.

P: I1(X) I1(Y) e1(X) I2(Y) e3(Y) e1(X) I2(Y)

¿SERIALIZABLE POR CONFLICTOS?



Conclusión:

Existe un ciclo, por lo que no es serializable por conflictos

¿SERIALIZABLE POR VISTAS?

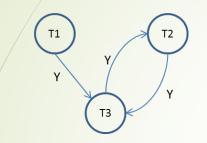
T1T2T3: I1(X) I1(Y) e1(X) e1(X) I2(Y) I2(Y) e3(Y) No es equivalente, dado que en el plan en serie la segunda lectura I2(Y) es limpia, mientras que en el plan a comparar es sucia.

T2T1T3: I2(Y) I2(Y) I1(X) I1(Y) e1(X) e1(X) e3(Y). No es equivalente, dado que en el plan en serie la segunda lectura I2(Y) es limpia, mientras que en el plan a comparar es sucia.

T3T1T2: e3(Y) I1(X) I1(Y) e1(X) e1(X) I2(Y) I2(Y) No es equivalente, dado que en el plan en serie las lecturas I1(Y) y la primera I2(Y) son sucias, mientras que en el plan a comparar son limpias.

P: I1(X) I1(Y) e1(X) I2(Y) e3(Y) e1(X) I2(Y)

¿SERIALIZABLE POR CONFLICTOS?



Conclusión:

Existe un ciclo, por lo que no es serializable por conflictos

¿SERIALIZABLE POR VISTAS?

T3T2T1: e3(Y) I2(Y) I2(Y) I1(X) I1(Y) e1(X) e1(X) No es equivalente, dado que en el plan en serie las lecturas I1(Y) y la primera I2(Y) son sucias, mientras que en el plan a comparar son limpias.

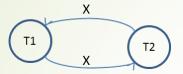
T1T3T2: I1(X) I1(Y) e1(X) e1(X) e3(Y) I2(Y) I2(Y) No es equivalente, dado que en el plan en serie la primera lectura I2(Y) es sucia, mientras que en el plan a comparar es limpia.

T2T3T1: I2(Y) I2(Y) e3(Y) I1(X) I1(Y) e1(X) e1(X). No es equivalente, dado que en el plan en serie la segunda lectura I2(Y) es limpia, mientras que en el plan a comparar es sucia.

Conclusión: **NO es serializable por vistas**, dado que no existe ningún plan en serie equivalente a P por vistas.

P: I1(X) e2(X) e1(X) a2 c1

¿SERIALIZABLE POR CONFLICTOS?



Conclusión:

Existe un ciclo, por lo que no es serializable por conflictos

¿SERIALIZABLE POR VISTAS?

T1 T2: I1(X) e1(X) c1 e2(X) a2 Es equivalente porque la lectura de X por parte de T1 es limpia en ambos planes. Además, la última escritura, e1(X), es la misma (dado que T2 se anula).

T2 T1: e2(X) a2 I1(X) e1(X) c1 Es equivalente porque la lectura de X por parte de T1 es limpia en ambos planes (dado que T2 se anula). Además, la última escritura, e1(X), es la misma.

Conclusión: Es serializable por vistas, teniendo en cuenta que T2 se anula.

Bibliografía

Fundamentos de Sistemas de Bases de Datos. Ramez A. Elmasri, Shamkant B. Navathe [cap. 19]

Fundamentos de Bases de Datos. A. Silberschatz. McGraw-Hill [cap. 15]