



Escuela Técnica Superior
de Ingeniería Informática
Departamento de Lenguajes
y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

*Grupo de Ingeniería del Software y Bases de Datos
Departamento de Lenguajes y Sistemas Informáticos
Universidad de Sevilla
septiembre 2016*



Escuela Técnica Superior
de Ingeniería Informática
Departamento de Lenguajes
y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

- **Objetivos de este tema**
 - Entender el concepto de **transacción** en bases de datos y sus propiedades deseables.
 - Conocer los problemas provocados por la **concurrencia** en bases de datos y sus posibles soluciones.
 - Conocer los **niveles de aislamiento** de transacciones en bases de datos SQL.
 - Conocer los distintos tipos de **fallos** en bases de datos y la forma de **recuperarse** de ellos.



septiembre 2016 Introducción a la Ingeniería del Software y a los Sistemas de Información 1



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

Gestión de Transacciones: Concurrencia y Recuperación

- ¿Qué es una transacción en una BD?
 - Es una **secuencia de operaciones** ($op_1..op_N$) que forma una **unidad lógica de trabajo** en una BD.
 - Normalmente, se corresponden con la ocurrencia de algún **hecho** en el entorno del sistema de información del que se deba tener memoria*.



*Recordar la función de memoria de los sistemas de información.

septiembre 2016

Introducción a la Ingeniería del Software y a los Sistemas de Información

2

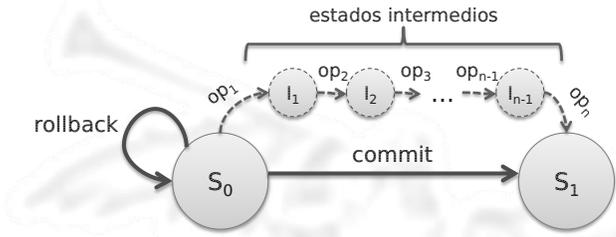


Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

Gestión de Transacciones: Concurrencia y Recuperación

- Propiedades deseables (ACID)
 - **Atómica**: si se confirma (**commit**), se ejecutan todas las operaciones; si se cancela (**rollback**), no se ejecuta ninguna.
 - No se puede ejecutar parcialmente, dejando la base de datos en un estado intermedio (I_i).



septiembre 2016

Introducción a la Ingeniería del Software y a los Sistemas de Información

3



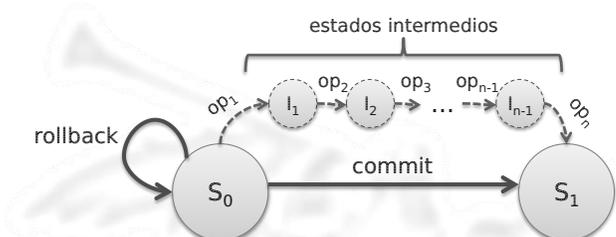
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

- **Propiedades deseables (ACID)**
 - **Consistente:** a partir de un estado inicial consistente (S_0), deja la base de datos en otro estado consistente (S_1), que respeta todas las **reglas de integridad**.
 - Los estados intermedios ($I_{1..n}$) no son necesariamente consistentes*.





*Por eso es necesario que sea atómica, para no dejar la base de datos en un estado inconsistente.

septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
4



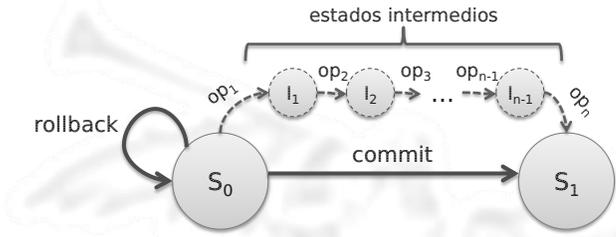
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

- **Propiedades deseables (ACID)**
 - **Aislada:** se realiza como si fuera la única transacción ejecutándose en la base de datos.
 - Los estados intermedios ($I_{1..n}$) no deben ser visibles para otras transacciones.
 - Es la propiedad **menos desarrollada** en los SGBD.





septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
5



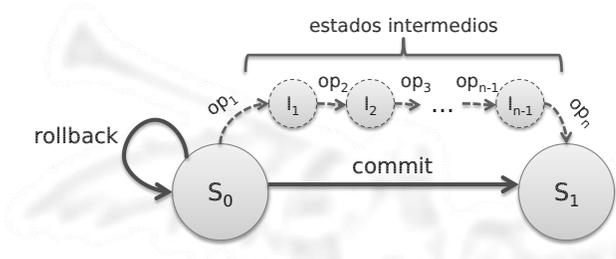
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

Gestión de Transacciones: Concurrencia y Recuperación

- **Propiedades deseables (ACID)**
 - **Durable**: si se confirma (**commit**), los cambios en la base de datos se hacen **permanentes**.
 - El SGBD es responsable de que los cambios sean permanentes, incluso aunque se produzca un **fallo**.





septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
6



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

Gestión de Transacciones: Concurrencia y Recuperación

- **Problemas de concurrencia**
 - **Múltiples usuarios simultáneos** del sistema de información provocan la ejecución de múltiples **transacciones concurrentes**.
 - Si el aislamiento no está garantizado por el SGBD, puede haber **problemas de consistencia**.



septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
7



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrency and Recovery

- **Actualización perdida (lost update)**
 - Unas transacciones sobrescriben las actualizaciones de otras.

Transacción 1	Transacción 2
$x_1 = \text{LEER}(X);$	
	$x_2 = \text{LEER}(X);$ $x_2 = x_2 + 1;$ ESCRIBIR(X, x_2); COMMIT;
$x_1 = x_1 + 1;$ ESCRIBIR(X, x_1); COMMIT;	

se pierde

septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
8



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrency and Recovery

- **Lectura sucia (dirty read)***
 - Se lee un valor que está siendo modificado por otra transacción que no ha finalizado y que podría cancelarse o fallar.

Transacción 1	Transacción 2
$x_1 = \text{LEER}(X);$ $x_1 = x_1 + 1;$ ESCRIBIR(X, x_1);	
	$x_2 = \text{LEER}(X);$ ESCRIBIR(X, x_2); COMMIT;
ROLLBACK;	

lectura sucia

*Denominado también *dependencia no confirmada*.

septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
9



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrency and Recovery

- **Lectura no repetible (non-repeatable read)**
 - Se obtienen lecturas diferentes del mismo valor durante la misma transacción.

Transacción 1	Transacción 2
$x_1 = \text{LEER}(X);$ IMPRIMIR(x_1);	
	$x_2 = \text{LEER}(X);$ $x_2 = x_2 + 1;$ ESCRIBIR(X, x_2); COMMIT;
$x_1 = \text{LEER}(X);$ IMPRIMIR(x_1); COMMIT	



septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
10



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrency and Recovery

- **Lectura fantasma (phantom read)**
 - Se obtienen lecturas diferentes de un conjunto de tuplas durante la misma transacción.

Transacción 1	Transacción 2
$\{x_i\} = \text{LEER}(\{X\});$ IMPRIMIR($\{x_i\});$	
	$\{x_j\} = \text{LEER}(\{X\});$ $\{x_j\} = \{x_j\} \cup \{x_k\};$ ESCRIBIR($\{X\}, \{x_j\});$ COMMIT;
$\{x_i\} = \text{LEER}(\{X\});$ IMPRIMIR($\{x_i\});$ COMMIT	



septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
11



Escuela Técnica Superior
de Ingeniería Informática
Departamento de Lenguajes
y Sistemas Informáticos

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

Gestión de Transacciones: Concurrencia y Recuperación

- ¿Cómo evitar los problemas de concurrencia?
 - Ejecutando las transacciones **secuencialmente**:
 - Reduce mucho el rendimiento de un SGBD.
$$T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow \dots \rightarrow T_n$$
 - Ejecutando las transacciones **concurrentemente** pero con un **plan de ejecución secuenciable**:
 - Debe tener el mismo efecto final que si las transacciones se ejecutaran secuencialmente (**secuenciabilidad**).
 - Normalmente, el SGBD usa **bloqueos**, marcas de tiempo (timestamps) o versiones múltiples de los datos para garantizar la secuenciabilidad.

12

septiembre 2016 Introducción a la Ingeniería del Software y a los Sistemas de Información



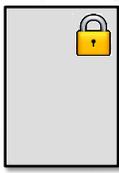
Escuela Técnica Superior
de Ingeniería Informática
Departamento de Lenguajes
y Sistemas Informáticos

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

Gestión de Transacciones: Concurrencia y Recuperación

- El concepto de gránulo de bloqueo
 - Un **gránulo** es una unidad que puede ser bloqueada por una transacción. 
 - Dependiendo del SGBD, el gránulo puede ser una tabla, un rango de filas, una sola fila, una columna de una fila, etc.
 - Si una transacción intenta bloquear un gránulo ya bloqueado, **espera** hasta su liberación.

Tabla



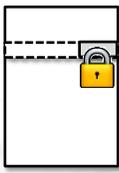
Rango



Fila



Columna



13

septiembre 2016 Introducción a la Ingeniería del Software y a los Sistemas de Información



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

Gestión de Transacciones: Concurrencia y Recuperación

- **Tipos de bloqueos**
 - **Para lectura** (compartido): un gránulo puede estar bloqueado para lectura por varias transacciones; los intentos de bloqueo exclusivo deben esperar.
 - **Para escritura** (exclusivo): un gránulo puede estar bloqueado para escritura por una sola transacción; los demás intentos de bloqueo deben esperar.

Tabla



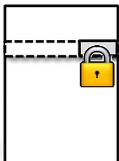
Rango



Fila



Columna





septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
14



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

Gestión de Transacciones: Concurrencia y Recuperación

- **Aislamiento de transacciones en SQL**
 - SQL-92 define cuatro niveles de aislamiento dentro de una transacción:
 - **READ_UNCOMMITTED**: permite leer datos no confirmados de otras transacciones.
 - **READ_COMMITTED**: sólo permite leer datos confirmados antes de que se ejecuta una consulta. Suele ser el valor por defecto.
 - **REPEATABLE_READ**: sólo permite leer datos confirmados antes de que comience la transacción.
 - **SERIALIZABLE**: garantiza la secuenciabilidad de la transacción con todas las demás transacciones concurrentes; si no puede garantizarlo, la transacción falla.

septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
15



UNIVERSIDAD DE SEVILLA
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

- Aislamiento de transacciones en SQL
 - Se debe especificar al comienzo de la transacción: **SET TRANSACTION ISOLATION LEVEL** nivel;
 - El nivel por defecto suele ser **READ_COMMITTED**, que ofrece un buen rendimiento.
 - Para solucionar el problema de la **actualización perdida** con nivel **READ_COMMITTED**, algunos SGBD permiten **bloquear para escritura** al hacer un **SELECT** mediante la cláusula **FOR UPDATE**:


```
SELECT * FROM tabla
                WHERE condición
                FOR UPDATE; ← bloquea para escritura
```

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

septiembre 2016 Introducción a la Ingeniería del Software y a los Sistemas de Información 16



UNIVERSIDAD DE SEVILLA
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

- Aislamiento de transacciones en SQL

Nivel de aislamiento	Problemas de concurrencia			
	Lost update	Dirty read	Non-repeat. read	Phantom read
READ_UNCOMMITTED	✗	✗	✗	✗
READ_COMMITTED	✗*	✓	✗	✗
REPEATABLE_READ	✓	✓	✓	✗
SERIALIZABLE	✓	✓	✓	✓

– A mayor nivel de aislamiento, menor **rendimiento**.

* Salvo que se use la cláusula FOR UPDATE (ver transparencia anterior).

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

septiembre 2016 Introducción a la Ingeniería del Software y a los Sistemas de Información 17



UNIVERSIDAD DE SEVILLA
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

- Aislamiento de transacciones en SQL

Nivel de aislamiento	Bloqueos		
	Bloqueo Escritura	Bloqueo Lectura	Bloqueo de Rango
READ_UNCOMMITTED			
READ_COMMITTED			
REPEATABLE_READ			
SERIALIZABLE			

-  Mantiene el bloqueo hasta el final de la transacción.
-  Mantiene el bloqueo hasta el final de la operación.

Fuente: https://es.wikipedia.org/wiki/Aislamiento_%28ACID%29

septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
18



UNIVERSIDAD DE SEVILLA
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

- Aislamiento de transacciones en SQL
 - ¿Cómo se comportan los escenarios de los problemas de concurrencia en cada nivel de aislamiento?
 - ¿Puede ocurrir un interbloqueo de dos transacciones?



septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
19



UNIVERSIDAD DE SEVILLA
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

1. El concepto de transacción

2. Problemas de la concurrencia

3. Bloqueos en bases de datos

4. Aislamiento en SQL

5. Recuperación ante fallos

- Tipos de fallos en bases de datos
 - Fallo de una orden al SGBD
 - Violación de reglas de integridad, tipos incorrectos, nombres de tablas o columnas erróneos, errores de sintaxis, división por cero, etc.
 - Lo detecta el SGBD y lo debe tratar la aplicación mediante el tratamiento de excepciones oportuno.
 - Fallo de transacción
 - Error de programación, interbloqueo, fallo de una orden al SGBD, etc.
 - Puede detectarlo tanto el SGBD como la aplicación, aunque el tratamiento (ROLLBACK) es responsabilidad del SGBD.



septiembre 2016

Introducción a la Ingeniería del Software y a los Sistemas de Información

20



UNIVERSIDAD DE SEVILLA
Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

1. El concepto de transacción

2. Problemas de la concurrencia

3. Bloqueos en bases de datos

4. Aislamiento en SQL

5. Recuperación ante fallos

- Tipos de fallos en bases de datos
 - Fallo del sistema
 - Fallo de hardware, problemas de alimentación eléctrica, etc. sin que el medio de almacenamiento se vea afectado.
 - El SGBD es responsable de su recuperación mediante el uso de bitácoras (**logs**).
 - Los logs siempre se escriben en disco **antes** que las transacciones, para poder recuperarlas después.





septiembre 2016

Introducción a la Ingeniería del Software y a los Sistemas de Información

21



Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

- **Tipos de fallos en bases de datos**
 - **Fallo del medio**
 - Fallo del hardware del medio de almacenamiento por avería, desastre natural, sabotaje, etc.
 - Fallo de programación que ha llevado a la pérdida permanente de datos al sobrescribirlos o borrarlos.
 - La recuperación es responsabilidad de los administradores de la base de datos, normalmente mediante copias de seguridad (**backups**).




septiembre 2016

Introducción a la Ingeniería del Software y a los Sistemas de Información

22

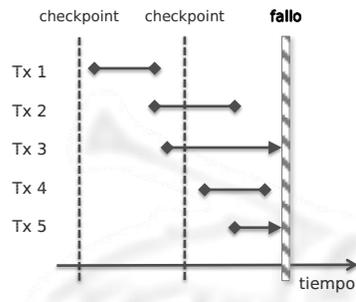


Escuela Técnica Superior de Ingeniería Informática
Departamento de Lenguajes y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

- **Algoritmos de recuperación de transacciones**
 - Cada cierto tiempo, el SGBD para el procesamiento de transacciones y vuelca todos los datos en memoria en el sistema de almacenamiento: **checkpoint**.



Tx 1		Tx 1: no hacer nada, está en disco.
Tx 2		Tx 2: rehacer, completa pero no en disco.
Tx 3		Tx 3: deshacer, incompleta.
Tx 4		Tx 4: rehacer, completa pero no en disco.
Tx 5		Tx 5: deshacer, incompleta.

septiembre 2016

Introducción a la Ingeniería del Software y a los Sistemas de Información

23



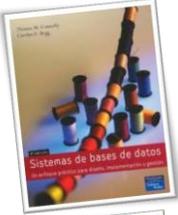
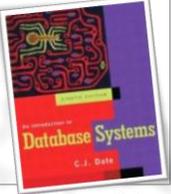
Escuela Técnica Superior
de Ingeniería Informática
Departamento de Lenguajes
y Sistemas Informáticos

Gestión de Transacciones: Concurrencia y Recuperación

1. El concepto de transacción
2. Problemas de la concurrencia
3. Bloqueos en bases de datos
4. Aislamiento en SQL
5. Recuperación ante fallos

• Bibliografía

- R. Elmasri, S. Navathe, **Fundamentos de Sistemas de Bases de Datos** (5ª edición). Ed. Addison-Wesley, 2007.
 - Capítulos 17, 18 y 19.
- T. Connolly, C. Begg, **Sistemas de Bases de Datos** (4ª edición). Ed. Addison-Wesley, 2005.
 - Capítulo 20.
- C. J. Date, **An Introduction to Database Systems** (8th edition). Ed. Addison-Wesley, 2003.
 - Capítulos 15 y 16.

septiembre 2016
Introducción a la Ingeniería del Software y a los Sistemas de Información
24