

# CSE 444: Database Internals

## Section 5: Transactions

# Today

- Serializability and Conflict Serializability
  - Precedence graph
- Two-Phase Locking
  - Strict two phase locking
- Timestamp-based Concurrency Control
- Multiversion Concurrency Control

# Problem 1: Serializability and Locking

- Is this schedule conflict serializab

What is

- Serializability
- Conflict Serializability?

$T_0$	$T_1$
$R_0(A)$	
$W_0(A)$	
	$R_1(A)$
	$R_1(B)$
	$C_1$
$R_0(B)$	
$W_0(B)$	
$C_0$	

# Review: (Conflict) Serializable Schedule

- A schedule is ***serializable*** if it is equivalent to a serial schedule
- A schedule is ***conflict serializable*** if it can be transformed into a serial schedule by a series of swappings of adjacent non-conflicting actions

# Review: (Conflict) Serializable Schedule

- A schedule is **serializable** if it is equivalent to a serial schedule
- A schedule is **conflict serializable** if it can be transformed into a serial schedule by a series of swappings of adjacent non-conflicting actions

Example:

$r_1(A); w_1(A); r_2(A); w_2(A); r_1(B); w_1(B); r_2(B); w_2(B)$

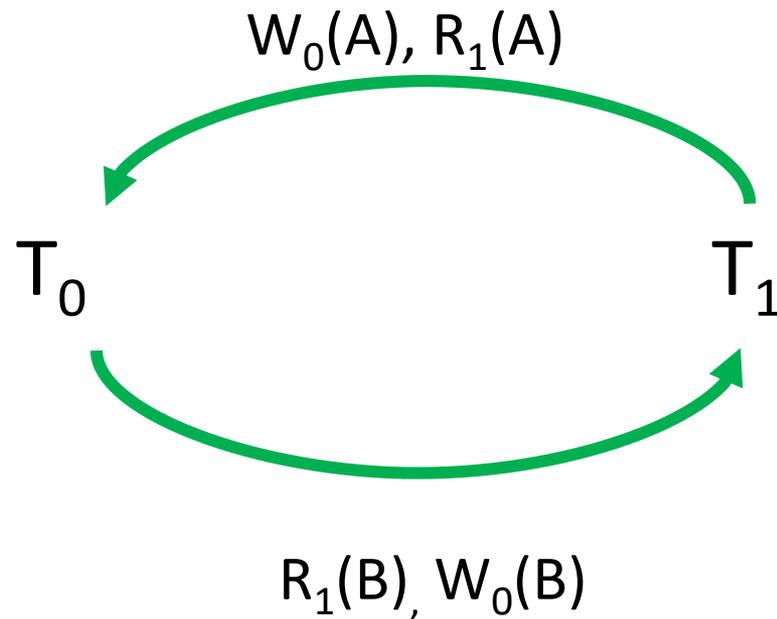
$r_1(A); w_1(A); r_1(B); w_1(B); r_2(A); w_2(A); r_2(B); w_2(B)$

# Problem 1: Serializability and Locking

- Is this schedule conflict serializable?

$T_0$	$T_1$
$R_0(A)$	
$W_0(A)$	
	$R_1(A)$
	$R_1(B)$
	$C_1$
$R_0(B)$	
$W_0(B)$	
$C_0$	

- No.
- The **precedence graph** contains a cycle



- So, use 2PL ...
  - Original schedule below

$T_0$	$T_1$
$R_0(A)$	
$W_0(A)$	
	$R_1(A)$
	$R_1(B)$
	$C_1$
$R_0(B)$	
$W_0(B)$	
$C_0$	

- So, use 2PL ...

❑ Original schedule below

What is

- Two Phase Locking
- Strict Two Phase Locking?

$T_0$	$T_1$
$R_0(A)$	
$W_0(A)$	
	$R_1(A)$
	$R_1(B)$
	$C_1$
$R_0(B)$	
$W_0(B)$	
$C_0$	

# Review:

## (Strict) Two Phase Locking (2PL)

### The 2PL rule:

In every transaction, all lock requests must precede all unlock requests

### Strict 2PL:

All locks held by a transaction are released when the transaction is completed

- Ensures that schedules are recoverable
  - Transactions commit only after all transactions whose changes they read also commit
- Avoids cascading rollbacks

- How can 2PL can ensure a conflict-serializable schedule?

□ Original schedule below

$T_0$	$T_1$
$R_0(A)$	
$W_0(A)$	
	$R_1(A)$
	$R_1(B)$
	$C_1$
$R_0(B)$	
$W_0(B)$	
$C_0$	







$T_0$	$T_1$
$L_0(A)$	
$R_0(A)$	
$W_0(A)$	
	<b><math>L_1(A) : \text{Block}</math></b>
$L_0(B)$	
$R_0(B)$	
$W_0(B)$	
$U_0(A)$	
$U_0(B)$	
$C_0$	
	<b><math>L_1(A) : \text{Granted}</math></b>
	$R_1(A)$
	$L_1(B)$
	$R_1(B)$
	<b><math>U_1(A)</math></b>
	<b><math>U_1(B)</math></b>
	$C_1$

$T_0$	$T_1$
$L_0(A)$	
$R_0(A)$	
$W_0(A)$	
	<b><math>L_1(A)</math> : Block</b>
$L_0(B)$	Is this strict 2PL?
$R_0(B)$	
$W_0(B)$	No, release locks after commit
$U_0(A)$	
$U_0(B)$	
$C_0$	
	<b><math>L_1(A)</math> : Granted</b>
	$R_1(A)$
	$L_1(B)$
	$R_1(B)$
	<b><math>U_1(A)</math></b>
	<b><math>U_1(B)</math></b>
	$C_1$

# Problem 2: Timestamp-based Concurrency Control

# Timestamp-based Concurrency Control

- Some transaction,  $T$ .
- Some element (tuple/page),  $X$ .
- $TS(T)$  - timestamp for transaction  $T$ 
  - Stays constant for all of  $T$ 's operations
- $WT(X)$  – latest write timestamp for  $X$ 
  - Set  $WT(X) = TS(T)$
- $RT(X)$  – latest read timestamp for  $X$ 
  - Set  $RT(X) = TS(T)$
- $C(X)$  –  $X$ 's value has been committed
  - 1 if true, 0 if not

# Timestamp-based Concurrency Control

- **Actions for transaction T**
  - **Grant** a read/write request for a transaction
  - **Abort** (in case T violates physical reality – late actions)
  - **Delay** (make the Grant or Abort decision later)
    - When writing, the change is always tentative until we decide to commit. For this, we use a commit bit C to keep track if the transaction that last wrote X has committed
  - **Ignore *Thomas Write Rule*** – ignore outdated writes

# Timestamp-based Concurrency Control - Four Rules

- **Rule 1: Read** request on  $X$  by  $T$ 
  - $TS(T) < WT(X)$ , **abort**, (read too late)
  - $TS(T) \geq WT(X)$ , physically realizable
    - If  $C = 1$ , **grant**, update  $RT(X)$
    - If  $C = 0$ , **delay**  $T$

# Timestamp-based Concurrency Control - Four Rules

- **Rule 2: Write** request on **X** by **T**
  - $TS(T) < RT(X)$  (write too late)
    - **Abort**
  - $TS(T) \geq RT(X)$ , physically realizable
    - $TS(T) \geq WT(X)$ 
      - then **grant**, update  $WT(X)$ , set  $C = 0$  (as it's not committed yet)
    - $TS(T) < WT(X)$ 
      - If  $C = 1$ , **ignore** (*Thomas Write Rule* – ignore outdated writes)
      - If  $C = 0$ , **delay**

# Timestamp-based Concurrency Control - Four Rules

- **Rule 3: Commit** request by **T**
  - Set  $C = 1$  for all **X** written by **T**
  - Allow waiting transactions to proceed
- **Rule 4: Abort** transaction **T**
  - Check if the waiting transactions can proceed now.

# Timestamp-based Concurrency Control

Two transactions get started.

- $\text{Start}(T_1) \rightarrow \text{Start}(T_2)$

# Timestamp-based Concurrency Control

What will happen at the last request?

- $\text{Start}(T_1) \rightarrow \text{Start}(T_2) \rightarrow R_{T_1}(A) \rightarrow R_{T_2}(A) \rightarrow W_{T_1}(B) \rightarrow \mathbf{W_{T_2}(B)}$

# Timestamp-based Concurrency Control

What will happen at the last request?

- $\text{Start}(T_1) \rightarrow \text{Start}(T_2) \rightarrow R_{T_1}(A) \rightarrow R_{T_2}(A) \rightarrow W_{T_1}(B) \rightarrow \mathbf{W_{T_2}(B)}$   
– **ACCEPTED** [no need to check C(B)]

# Timestamp-based Concurrency Control

What will happen at the last request?

- $\text{Start}(T_1) \rightarrow \text{Start}(T_2) \rightarrow R_{T_1}(A) \rightarrow R_{T_2}(A) \rightarrow W_{T_1}(B) \rightarrow \mathbf{W_{T_2}(B)}$   
– **ACCEPTED** [no need to check C(B)]
- $\text{Start}(T_1) \rightarrow \text{Start}(T_2) \rightarrow R_{T_2}(A) \rightarrow \text{Commit}_{T_2} \rightarrow R_{T_1}(A) \rightarrow \mathbf{W_{T_1}(A)}$

# Timestamp-based Concurrency Control

## What will happen at the last request?

- $\text{Start}(T_1) \rightarrow \text{Start}(T_2) \rightarrow R_{T_1}(A) \rightarrow R_{T_2}(A) \rightarrow W_{T_1}(B) \rightarrow \mathbf{W_{T_2}(B)}$   
– **ACCEPTED** [no need to check C(B)]
  
- $\text{Start}(T_1) \rightarrow \text{Start}(T_2) \rightarrow R_{T_2}(A) \rightarrow \text{Commit}_{T_2} \rightarrow R_{T_1}(A) \rightarrow \mathbf{W_{T_1}(A)}$   
– **ABORT**  $T_1$  because  $R_{T_2}(A)$  precedes

# Problem 2: Timestamp-based Concurrency Control

- $TS_1 \rightarrow TS_2 \rightarrow TS_3 \rightarrow TS_4 \rightarrow R_1(X) \rightarrow R_2(X) \rightarrow$   
 $W_2(X) \rightarrow W_1(X) \rightarrow W_3(Y) \rightarrow W_2(Y) \rightarrow C_3 \rightarrow$   
 $W_4(Z) \rightarrow C_4 \rightarrow R_2(Z)$
- Remember!
  - Note changes to RT, WT, A and C bit for each element
  - Apply four rules

























$ST_1 \rightarrow ST_2 \rightarrow ST_3 \rightarrow ST_4 \rightarrow R_1(X) \rightarrow R_2(X) \rightarrow W_2(X) \rightarrow W_1(X) \rightarrow W_3(Y) \rightarrow W_2(Y) \rightarrow C_3 \rightarrow W_4(Z) \rightarrow C_4 \rightarrow R_2(Z)$

T1	T2	T3	T4	X	Y	Z
1	2	3	4	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1
$R_1(X)$				RT=1		
	$R_2(X)$			RT=2		
	$W_2(X)$			WT=2, C=0		
$W_1(X)$ : abort						
		$W_3(Y)$			WT=3, C=0	
	$W_2(Y)$ : delay					

1. Physically realizable:  
 $TS(T_3) \geq RT(X)$  although  $TS(T_2) < WT(X)$

2. We could not apply Thomas' write rule (**ignore  $W_2(Y)$** ) since  $C=0$





$ST_1 \rightarrow ST_2 \rightarrow ST_3 \rightarrow ST_4 \rightarrow R_1(X) \rightarrow R_2(X) \rightarrow W_2(X) \rightarrow W_1(X) \rightarrow W_3(Y) \rightarrow W_2(Y) \rightarrow C_3 \rightarrow W_4(Z) \rightarrow C_4 \rightarrow R_2(Z)$

T1	T2	T3	T4	X	Y	Z
1	2	3	4	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1
$R_1(X)$				RT=1		
	$R_2(X)$			RT=2		
	$W_2(X)$			WT=2, C=0		
$W_1(X)$ : abort						
		$W_3(Y)$			WT=3, C=0	
	$W_2(Y)$ : delay					
		$C_3$			C=1	

A later write by T<sub>3</sub> has been committed!

$ST_1 \rightarrow ST_2 \rightarrow ST_3 \rightarrow ST_4 \rightarrow R_1(X) \rightarrow R_2(X) \rightarrow W_2(X) \rightarrow W_1(X) \rightarrow W_3(Y) \rightarrow W_2(Y) \rightarrow C_3 \rightarrow W_4(Z) \rightarrow C_4 \rightarrow R_2(Z)$

T1	T2	T3	T4	X	Y	Z
1	2	3	4	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1
$R_1(X)$				RT=1		
	$R_2(X)$			RT=2		
	$W_2(X)$			WT=2, C=0		
$W_1(X)$ : abort						
		$W_3(Y)$			WT=3, C=0	
	$W_2(Y)$ : delay					
		$C_3$			C=1	
	<b>Ignore <math>W_2(Y)</math> and proceed</b>					

$ST_1 \rightarrow ST_2 \rightarrow ST_3 \rightarrow ST_4 \rightarrow R_1(X) \rightarrow R_2(X) \rightarrow W_2(X) \rightarrow W_1(X) \rightarrow W_3(Y) \rightarrow W_2(Y) \rightarrow C_3 \rightarrow W_4(Z) \rightarrow C_4 \rightarrow R_2(Z)$

T1	T2	T3	T4	X	Y	Z
1	2	3	4	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1
	Ignore $W_2(Y)$ and <b>proceed</b>					
			$W_4(Z)$			

$ST_1 \rightarrow ST_2 \rightarrow ST_3 \rightarrow ST_4 \rightarrow R_1(X) \rightarrow R_2(X) \rightarrow W_3(X) \rightarrow W_1(X) \rightarrow W_2(Y) \rightarrow W_2(Y) \rightarrow C_3 \rightarrow W_4(Z) \rightarrow C_4 \rightarrow$

1. Physically realizable:

$TS(T_4) \geq RT(X)$  and  $TS(T_4) \geq WT(X)$

2. Update WT and C (not committed yet)

Y	Z
RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1
	WT=4, C = 0

ignore  $w_2(Y)$   
and **proceed**

$W_4(Z)$

$ST_1 \rightarrow ST_2 \rightarrow ST_3 \rightarrow ST_4 \rightarrow R_1(X) \rightarrow R_2(X) \rightarrow W_2(X) \rightarrow W_1(X) \rightarrow W_3(Y) \rightarrow W_2(Y) \rightarrow C_3 \rightarrow W_4(Z) \rightarrow C_4 \rightarrow R_2(Z)$

T1	T2	T3	T4	X	Y	Z
1	2	3	4	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1
	Ignore $W_2(Y)$ and <b>proceed</b>					
			$W_4(Z)$			WT=4, C = 0
			$C_4$			C=1

$ST_1 \rightarrow ST_2 \rightarrow ST_3 \rightarrow ST_4 \rightarrow R_1(X) \rightarrow R_2(X) \rightarrow W_2(X) \rightarrow W_1(X) \rightarrow W_3(Y) \rightarrow W_2(Y) \rightarrow C_3 \rightarrow W_4(Z) \rightarrow C_4 \rightarrow R_2(Z)$

T1	T2	T3	T4	X	Y	Z
1	2	3	4	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1
	Ignore $W_2(Y)$ and <b>proceed</b>					
			$W_4(Z)$			WT=4, C = 0
			$C_4$			C=1
	$R_2(Z)$					

ST<sub>1</sub> -> ST<sub>2</sub> -> ST<sub>3</sub> -> ST<sub>4</sub> -> R<sub>1</sub>(X) -> R<sub>2</sub>(X) -> W<sub>2</sub>(X) -> W<sub>1</sub>(X) -> W<sub>3</sub>(Y) -> W<sub>2</sub>(Y) -> C<sub>3</sub> -> W<sub>4</sub>(Z) -> C<sub>4</sub> -> R<sub>2</sub>(Z)

1. **NOT** Physically realizable:

TS(T<sub>2</sub>) < WT(Z)

Abort/rollback

and proceed

R<sub>2</sub>(Z): abort

T4	X	Y	Z
4	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1	RT = 0, WT = 0, C = 1
W <sub>4</sub> (Z)			WT=4, C = 0
C <sub>4</sub>			C=1

# Timestamp-based Concurrency Control

Questions?

# Multiversion Concurrency Control

- Maintains **old** versions of database elements in addition the current version in the database itself.
- The idea is to allow reads that would otherwise result in an abort (as the current version was written by future transaction)

# Problem with Timestamp-Based Scheduling

T1	T2	T3	T4	A
150	200	175	225	RT = 0 WT = 0
R <sub>1</sub> (A)				RT = 150
W <sub>1</sub> (A)				WT = 150
	R <sub>2</sub> (A)			RT = 200
	W <sub>2</sub> (A)			WT = 200
		R <sub>3</sub> (A)		
		<b>Abort</b>		
			R <sub>4</sub> (A)	RT = 225

Had to abort because WT(A) is greater than my own timestamp

Would have been useful if I had access to an old version of A (from 150)...

# Multiversion Timestamps

T1	T2	T3	T4	A <sub>0</sub>	A <sub>150</sub>	A <sub>225</sub>
150	200	175	225	RT = 0 WT = 0		
R <sub>1</sub> (A)				Read		
W <sub>1</sub> (A)					Create	
	R <sub>2</sub> (A)				Read	
	W <sub>2</sub> (A)					Create
		R <sub>3</sub> (A)			Read	
			R <sub>4</sub> (A)			Read

Don't have to abort

Just read a previous value of  
A