

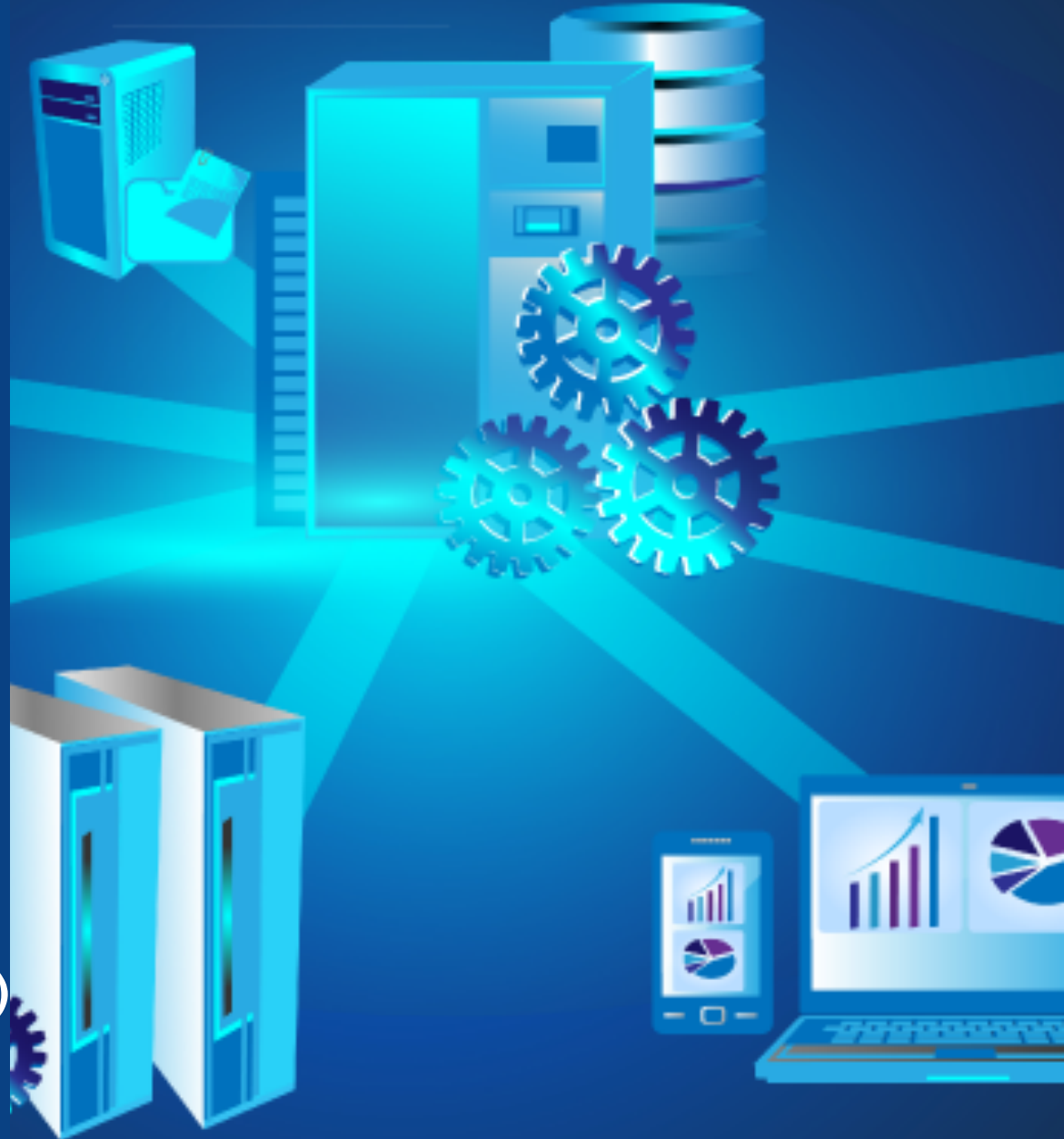


THE UNIVERSITY OF
MELBOURNE

**Semester 2,
2024**

Lecturer: Farhana Choudhury (PhD)
Live lecture – Week 10

COMP90050 Advanced Database Systems





Today's topics

- Crash recovery summary
- ARIES example
- Q/A
 - Group presentations from next week
 - Group report submission in Week 12
 - Final exam

We have seen crash recovery

What needs to be recovered if a crash happens?

- Has it been made durable - good!
- If not durable – what additional information are needed to recover them?

Data pages in the buffer

Crash manager maintains both durability and atomicity

在data page上的有的是committed transaction 完成的 有的不是 (要确保是committed transaction 完成的 那些东西 durable)

The changes by committed transactions – make them durable
The changes by aborted/running transactions - undo



What additional information we need to recover from a crash

Dirty Page table (WAL in place)

Pag e#	Oldest LSN (least Recent LSN)

X-table

Xid	Statu s	Last LSN

Log

Prev LSN	Tid	Type	P ag e	Le ng th	Of fs et	Old Value	New Value

Checkpoints: Periodically, the DBMS creates a checkpoint with current *Xact table* and *dirty page table*

Store logs and checkpoint records in a safe place



ARIES example: After a crash, we found these logs

Those actions are undone. (只要有这个transaction的action或changes都 undone).

- 0 BEGIN CHECKPOINT
- 5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
- 10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
- 15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
- 20 T1: COMMIT



Recovery: The Analysis Phase

X-table

Xid	Status	Last LSN
T1	R	
T2	R	

- ✓ Reconstruct state at checkpoint.
 - via **end_checkpoint** record.
- ✓ Scan log forward from checkpoint.
 - **End** record: Remove Xact from Xact table.
 - **Other records**: Add Xact to Xact table, set **lastLSN=LSN**, change Xact status on **commit**.
 - **Update** record: If P not in Dirty Page Table,
 - u Add P to D.P.T., set its **recLSN=LSN**.

Dirty Page table

Page #	Oldest LSN (Recent LSN)
P5	
P6	
P7	



LSN log sequence

ARIES example

LSN

- 0 BEGIN CHECKPOINT
5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
20 T1: COMMIT

① 写的 如果这个地有结 transaction id 就进 X-table
7 DPT

② (P1 进 DPT, P1 进) (先看 DPT 再 X-table)

③ No entry for P₂, P₂ in

★ 先 DPT 再 X-table ★

X-table

Xid	Status	Last LSN
② T ₁	R→C	② 10 ③ 15

Dirty Page table

Page #	Oldest LSN (Recent LSN)
② P ₁	② 10
③ P ₂	③ 15

Analysis phase:

LSN 5: Initialize XACT table and DPT to empty.

What is the Last LSN update changed by T₁
first LSN that made the Page dirty


```

0    BEGIN CHECKPOINT
5    END CHECKPOINT (EMPTY XACT TABLE AND DPT)
10   T1: UPDATE P1 (OLD: YYY NEW: ZZZ) ①
15   T1: UPDATE P2 (OLD: WWW NEW: XXX)
20   T1: COMMIT

```

Dirty Page

Page	Oldest LSN
..	

X-table
LSN: 10是第一个弄脏P的LSN, 即使后来有
10完了, 在LSN=15时, 不再更新old LSN

dirty page table

X-table

Xid	Status	Last LSN
① T ₁	① R	① 10
② T ₂	② A	② 15

page #	oldest LSN (Recent LSN)
① P ₁	① 10

所以
15不是最新
10 → 15X
10 remain the same

Analysis后得到 X-tab 与 DPT



ReDO



ARIES example

- 0 BEGIN CHECKPOINT
- 5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
- 10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
- 15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
- 20 T1: COMMIT

Dirty Page table

X-table

Xid	Status	Last LSN
T1	R	10

Page #	Oldest LSN (Recent LSN)
P1	10

Analysis phase:

LSN 5: Initialize XACT table and DPT to empty.

LSN 10: Add (T1, LSN 10) to XACT table. Add (P1, LSN 10) to DPT.



ARIES example

- 0 BEGIN CHECKPOINT
- 5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
- 10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
- 15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
- 20 T1: COMMIT

Dirty Page table

X-table

Xid	Status	Last LSN
T1	R	15

Page #	Oldest LSN (Recent LSN)
P1	10
P2	15

Analysis phase:

LSN 5: Initialize XACT table and DPT to empty.

LSN 10: Add (T1, LSN 10) to XACT table. Add (P1, LSN 10) to DPT.

LSN 15: Set LastLSN=15 for T1 in XACT table. Add (P2, LSN 15) to DPT.



ARIES example

- 0 BEGIN CHECKPOINT
- 5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
- 10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
- 15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
- 20 T1: COMMIT

Dirty Page table

X-table

Xid	Status	Last LSN
T1	C	15

Page #	Oldest LSN (Recent LSN)
P1	10
P2	15

Analysis phase:

LSN 5: Initialize XACT table and DPT to empty.

LSN 10: Add (T1, LSN 10) to XACT table. Add (P1, LSN 10) to DPT.

LSN 15: Set LastLSN=15 for T1 in XACT table. Add (P2, LSN 15) to DPT.

Recovery: The REDO Phase

- v We *repeat History* to reconstruct state at crash:
 - Reapply *all* updates (even of aborted Xacts!), redo CLR's.
- v Scan forward from log rec containing smallest *recLSN* in D.P.T. For each CLR or update log rec *LSN*, REDO the action unless:
 - Affected page is not in the Dirty Page Table, or
 - Affected page is in D.P.T., but has $\text{recLSN} > \text{LSN}$, or \times
 - *pageLSN* (in DB) $\nless \text{LSN}$.
- v To REDO an action:
 - Reapply logged action.
 - Set *pageLSN* to *LSN*. No additional logging!



ARIES example

Redo 从 DPT 中最小的 LSN 开始: 10 (开始) ①②③

Redo 看好 DPT

- 0 BEGIN CHECKPOINT
- 5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
- 10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
- 15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
- 20 T1: COMMIT

Dirty Page table

Page #	Oldest LSN (Recent LSN)
P1	10
P2	15

X-table

Xid	Status	Last LSN
T1	C	15

Why in ① check PageLSN store in the page?
 PageLSN is last LSN for which the page was written to disk, so all LSN before Last LSN of the page (i.e., all changes are already in that page have been

Redo phase: in the disk)

Scan forward through the log starting at LSN 10.

- ① LSN 10: Read page P1, check PageLSN stored in the page. If PageLSN < 10, redo LSN 10 (set value to ZZZ) and set the page's PageLSN = 10.

为什么 10 还没做 (not made durable yet) → 做 10 → ZZZ, page redo set LSN = 10

★ extra question in exam. ★

e.g. page P_1 's LSN has been found as 15
in database after crash \Rightarrow What would the redo
happens? $15 \neq 10 \Rightarrow$ We don't make any changes.
(不改 YYY 为 ZZZ)

P_1 处理完了, 看 P_2

② Check P_2 的 PageLSN, if $\text{pageLSN} < 15$, redo LSN 15
(set value to XXX) and set the page's PageLSN
11
15

Undo Only for the transaction that doesn't commit. \rightarrow

```
0 BEGIN CHECKPOINT
5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
20 T1: COMMIT
```

这里改为 T_2

X-table			Dirty Page	
Xid	Stat	Act	Page #	Old LSN (Re)

这个 T_1 commit 3 \Rightarrow 没有 failed transaction.

这个 undo 啥也不干.

Then undo T_2 , return the old value to P_2
choose the largest LSN among them, chain actions all undo.



ARIES example

- 0 BEGIN CHECKPOINT
- 5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
- 10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
- 15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
- 20 T1: COMMIT

Dirty Page table

X-table

Xid	Status	Last LSN
T1	C	15

Page #	Oldest LSN (Recent LSN)
P1	10
P2	15

Redo phase:

Scan forward through the log starting at LSN 10.

LSN 10: Read page P1, check PageLSN stored in the page. If $\text{PageLSN} < 10$, redo LSN 10 (set value to ZZZ) and set the page's $\text{PageLSN} = 10$.

LSN 15: Read page P2, check PageLSN stored in the page. If $\text{PageLSN} < 15$, redo LSN 15 (set value to XXX) and set the page's



ARIES example

- 0 BEGIN CHECKPOINT
- 5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
- 10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
- 15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
- 20 T1: COMMIT

Dirty Page table

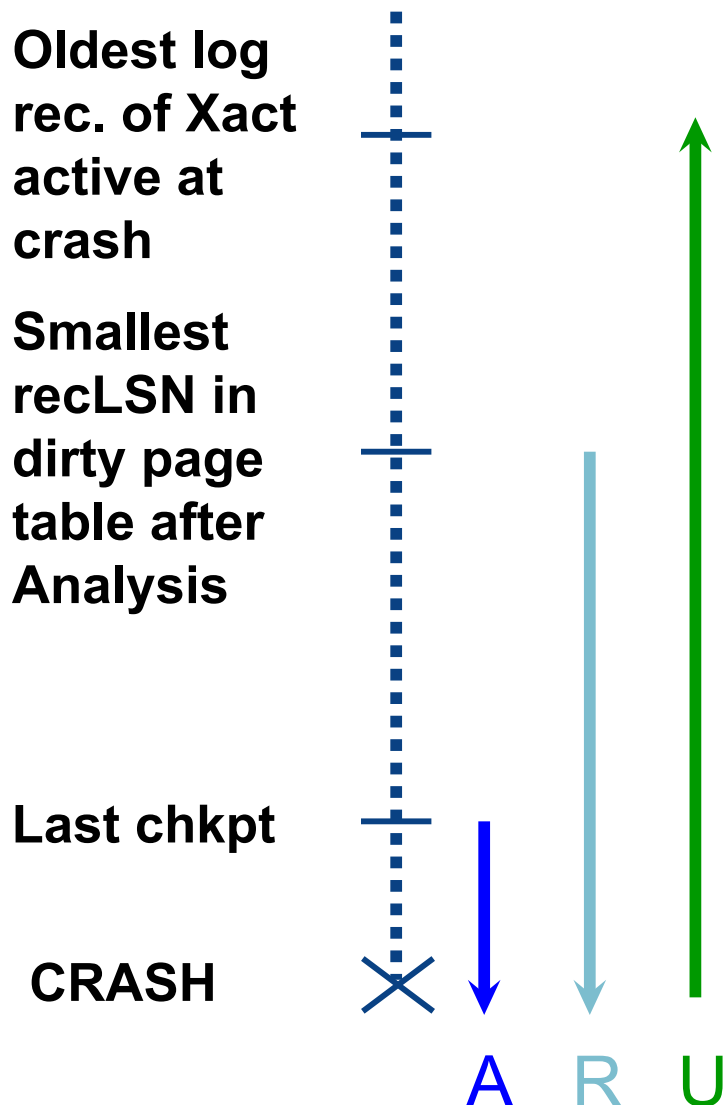
X-table

Xid	Status	Last LSN
T1	C	15

Page #	Oldest LSN (Recent LSN)
P1	10
P2	15

Undo phase:

Crash Recovery: Big Picture



- v Start from a **checkpoint** (found via **master** record).
- v Three phases. Need to:
 - Figure out which Xacts committed since checkpoint, which failed (**Analysis**).
 - **REDO** *all* actions.
 - u (repeat history)
 - **UNDO** effects of failed Xacts.



ARIES example

- 0 BEGIN CHECKPOINT
- 5 END CHECKPOINT (EMPTY XACT TABLE AND DPT)
- 10 T1: UPDATE P1 (OLD: YYY NEW: ZZZ)
- 15 T1: UPDATE P2 (OLD: WWW NEW: XXX)
- 20 T1: COMMIT

Dirty Page table

X-table

Xid	Status	Last LSN
T1	C	15

Page #	Oldest LSN (Recent LSN)
P1	10
P2	15

Undo phase:

Do nothing; no transactions to undo.



Q/A

- Group presentations from next week
- Group report submission in Week 12
- A sample final exam shared in canvas (the actual final exam will have more questions in it).