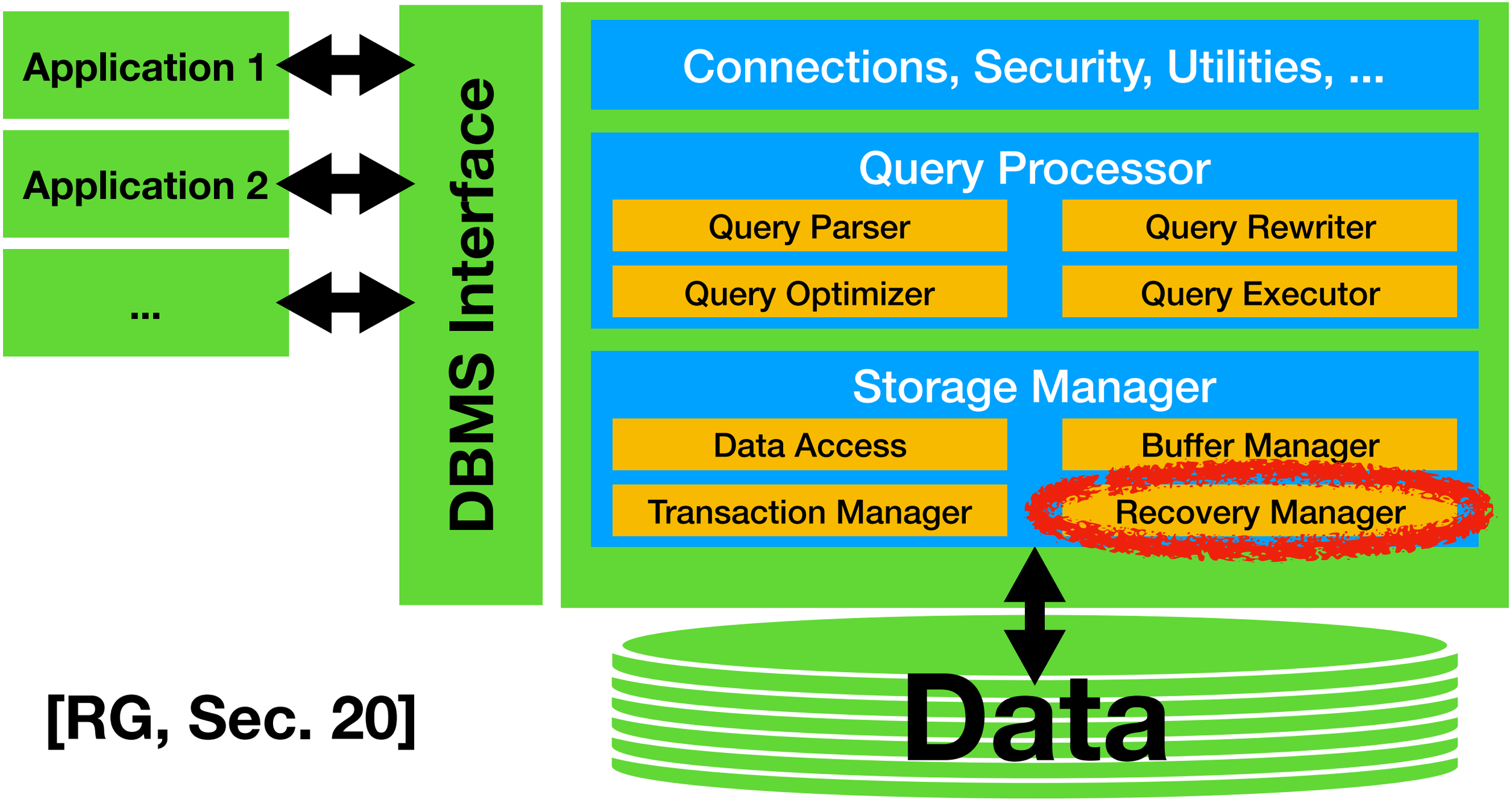
Recovery After System Crashes

Immanuel Trummer
itrummer@cornell.edu
www.itrummer.org

Database Management Systems (DBMS)



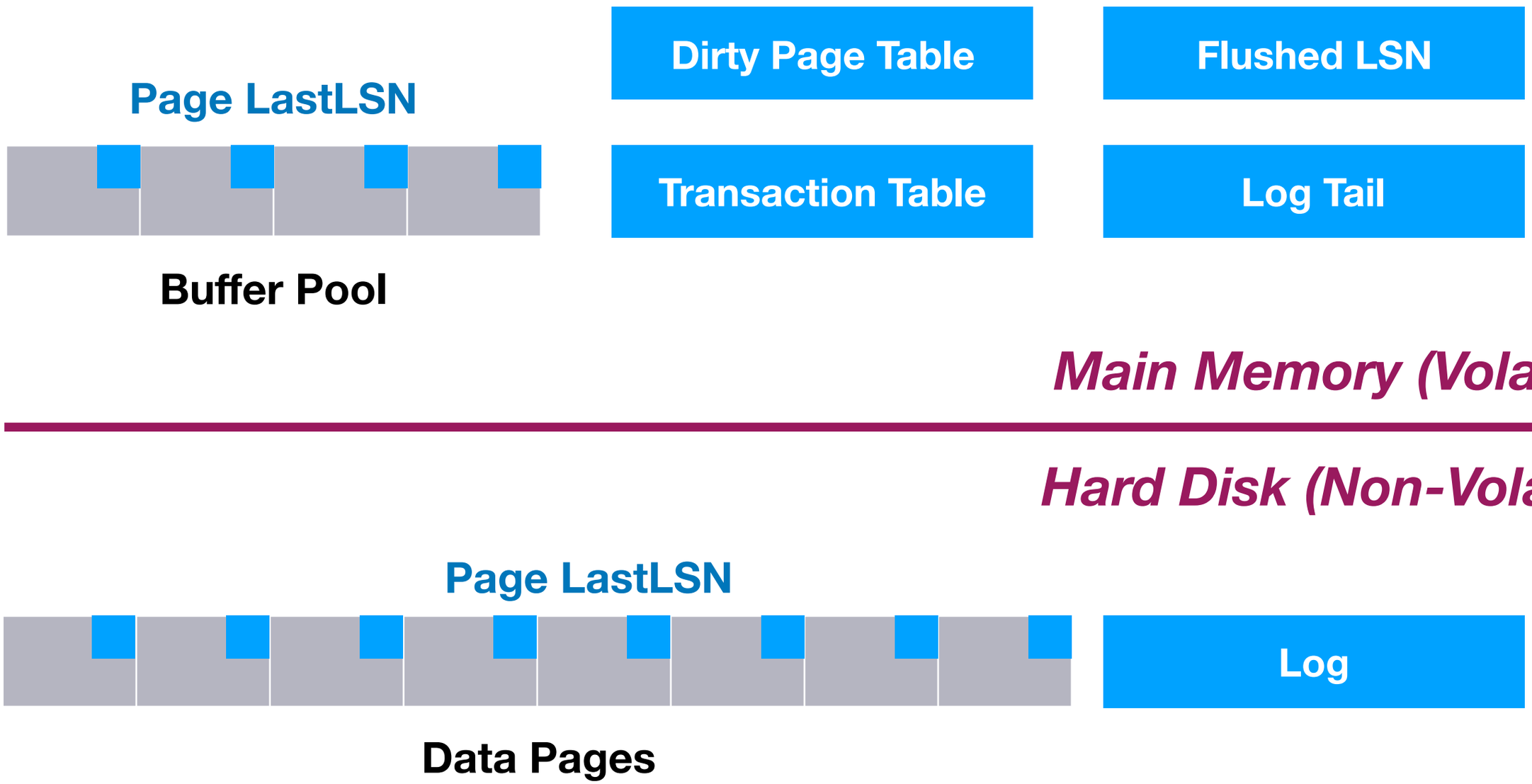
[RG, Sec. 20]

Slides by Immanuel Trummer, Cornell University

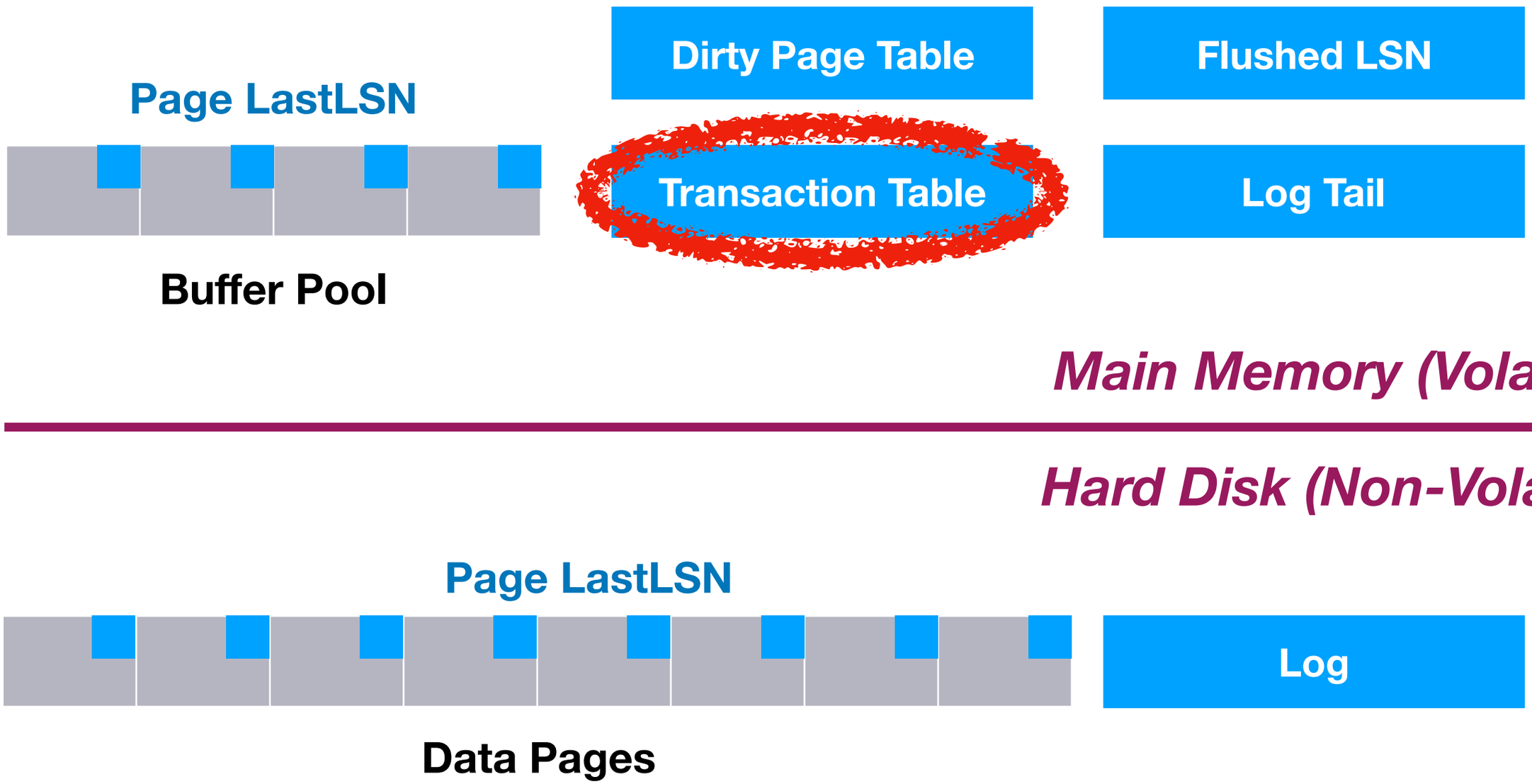
Outlook

- ARIES data structures
- **ARIES run time behavior**
- ARIES recovery algorithm

ARIES Data Structures



ARIES Data Structures



Updating Transaction Table

Scenario	Table Update
Transaction updates data	Update transaction lastLSN
Transaction commits	Update transaction status to committed
Transaction aborts	Update transaction status to aborted
Transaction ends	Remove from transaction table

ARIES Data Structures

Updating Dirty Page Table

Scenario	Table Update
Page Changed	If page not in table: add page with current LSN as recLSN
Page Written to Disk	Remove page if in table

ARIES Data Structures

Buffer Pool

Dirty Page Table

Flushed LSN

Transaction Table

Log Tail

Main Memory (Volatile)

Hard Disk (Non-Volatile)

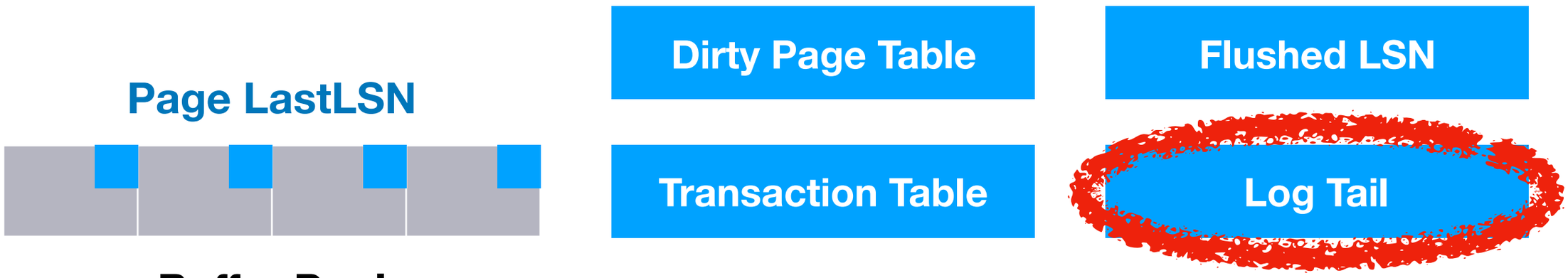
Data Pages

Log

Updating PageLSN

Scenario	Table Update
Data update	Update PageLSN in memory
Page Written to Disk	PageLSN is copied to disk

ARIES Data Structures



Updating Log Tail Buffer

Scenario	Log Update
Transaction updates data	Write update log entry
Transaction commits	Write commit log entry
Transaction aborts	Write abort log entry
Undo transaction update	Write compensation log record
Finished transaction cleanup	Write end log record

ARIES Data Structures

Dirty Page Table

Flushed LSN

Transaction Table

Log Tail

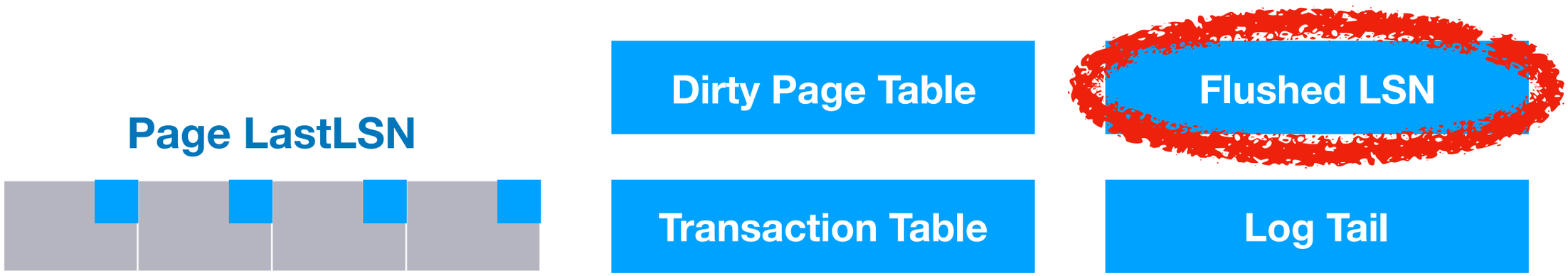
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Updating Log on Disk

Scenario	Disk Log Update
Transaction commits	Before commit: Flush log entries up to last transaction entry
Page written to disk	Before writing: Flush log entries until last entry affecting page (pageLSN)

ARIES Data Structures



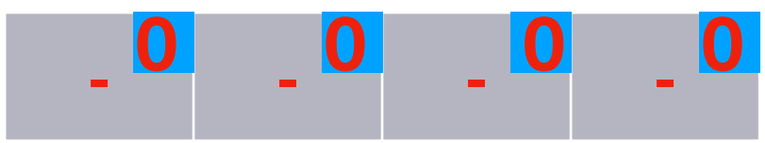
Updating FlushedLSN

Scenario	Update to FlushedLSN
Log written to hard disk until LSN X	Update FlushedLSN to X

ARIES Example (Run Time)

T-Table	DP Table	Log Tail

Page LastLSN



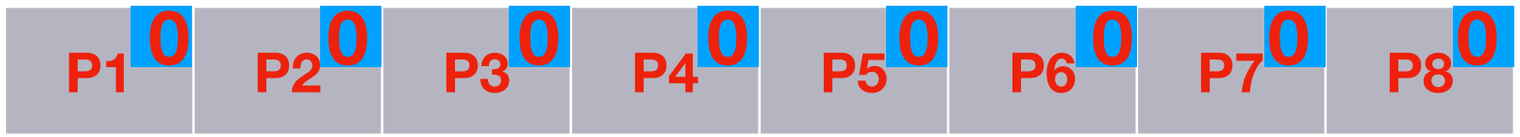
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

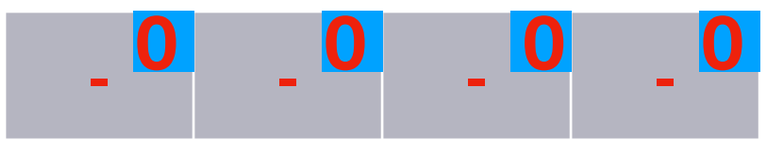


Data Pages

Log

ARIES Example (Run Time)

Page LastLSN



Buffer Pool

T-Table

DP Table

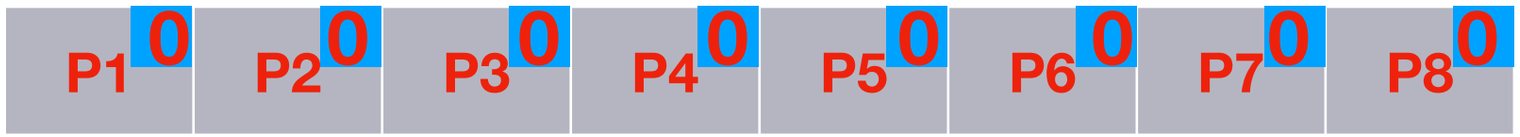
Log Tail
1: T1 Updates P2

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log

ARIES Example (Run Time)

Page LastLSN

P2	-	-	-
1	0	0	0

Buffer Pool

T-Table
T1, Running, 1

DP Table
P2, 1

Log Tail
1: T1 Updates P2

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1	P2	P3	P4	P5	P6	P7	P8
0	0	0	0	0	0	0	0

Data Pages

Log

ARIES Example (Run Time)

Page LastLSN

P2	-	-	-
1	0	0	0

Buffer Pool

T-Table
T1, Running, 1

DP Table
P2, 1

Log Tail
1: T1 Updates P2
2: T2 Updates P5

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1	P2	P3	P4	P5	P6	P7	P8
0	0	0	0	0	0	0	0

Data Pages

Log

ARIES Example (Run Time)

Page LastLSN

P2 1	P5 2	- 0	- 0

Buffer Pool

T-Table
T1, Running, 1
T2, Running, 2

DP Table
P2, 1
P5, 2

Log Tail
1: T1 Updates P2
2: T2 Updates P5

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log

ARIES Example (Run Time)

Page LastLSN

P2 1	P5 2	- 0	- 0

Buffer Pool

T-Table
T1, Running, 1
T2, Running, 2

DP Table
P2, 1
P5, 2

Log Tail
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log

ARIES Example (Run Time)

Page LastLSN

P2 3	P5 2	- 0	- 0

Buffer Pool

T-Table
T1, Running, 3
T2, Running, 2

DP Table
P2, 1
P5, 2

Log Tail
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log

ARIES Example (Run Time)

Page LastLSN

P2 3	P5 2	- 0	- 0

Buffer Pool

T-Table
T1, Running, 3
T2, Running, 2

DP Table
P2, 1
P5, 2

Log Tail
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log

ARIES Example (Run Time)

Page LastLSN

P2 3	P5 2	- 0	- 0

Buffer Pool

T-Table
T1, Running, 3
T2, Committed, 4

DP Table
P2, 1
P5, 2

Log Tail
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log

ARIES Example (Run Time)

T-Table
T1, Running, 3

DP Table
P2, 1
P5, 2

Log Tail
5: T2 End

Page LastLSN

P2 3	P5 2	- 0	- 0

Buffer Pool

FlushedLSN: 4

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

ARIES Example (Run Time)

Page LastLSN

P2 3	P5 2	- 0	- 0

Buffer Pool

T-Table
T1, Running, 3

DP Table
P2, 1
P5, 2

Log Tail
5: T2 End
6: T3 Updates P7

FlushedLSN: 4

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

ARIES Example (Run Time)

Page LastLSN

P2 3	P5 2	P7 6	- 0

Buffer Pool

T-Table
T1, Running, 3
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6

Log Tail
5: T2 End
6: T3 Updates P7

FlushedLSN: 4

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

ARIES Example (Run Time)

Page LastLSN

P2 3	P5 2	P7 6	- 0

Buffer Pool

T-Table
T1, Running, 3
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6

Log Tail
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

FlushedLSN: 4

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 0

Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

ARIES Example (Run Time)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

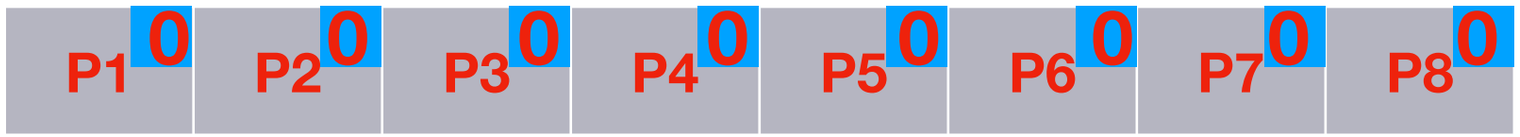
Log Tail
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

FlushedLSN: 4

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

ARIES Example (Run Time)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

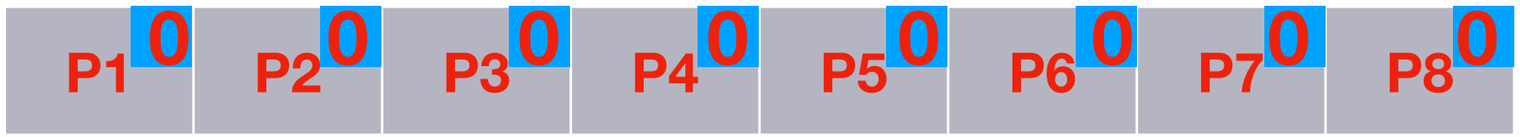
Log Tail
5: T2 End
6: T3 Updates P7
7: T1 Updates P8
8: T1 Updates P6

FlushedLSN: 4

Main Memory (Volatile)

Hard Disk (Non-Volatile)

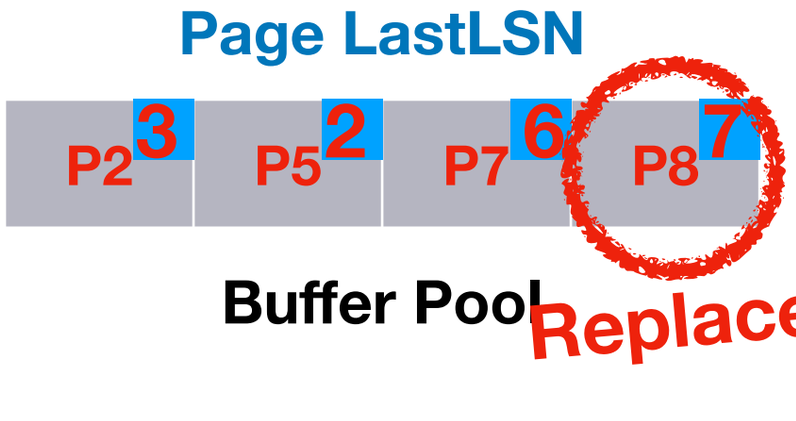
Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

ARIES Example (Run Time)



T-Table
T1, Running, 7
T3, Running, 6

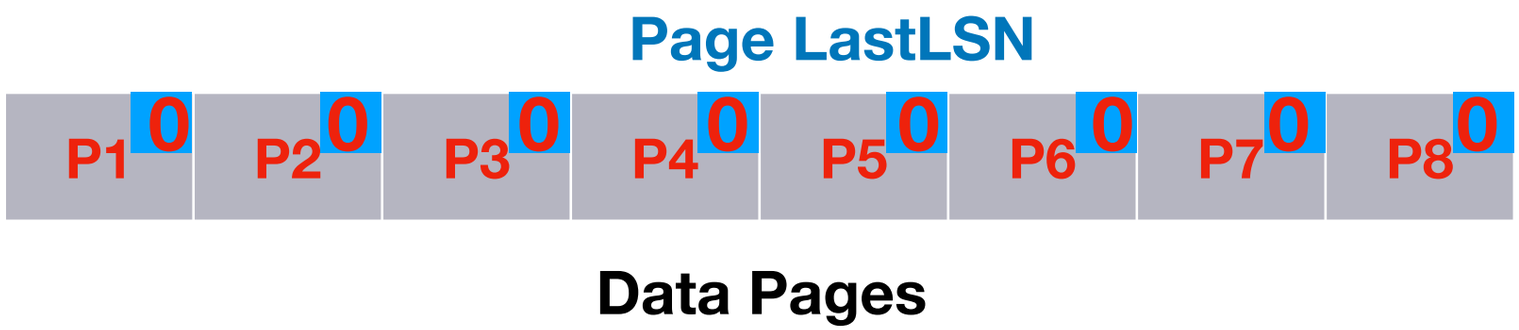
DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail
5: T2 End
6: T3 Updates P7
7: T1 Updates P8
8: T1 Updates P6

FlushedLSN: 4

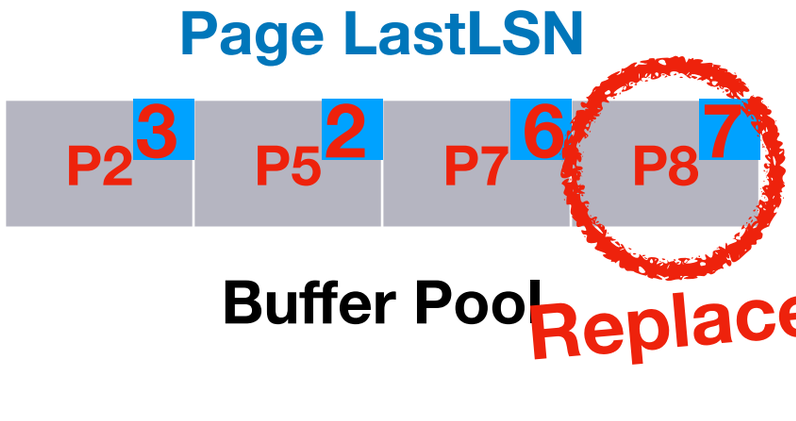
Main Memory (Volatile)

Hard Disk (Non-Volatile)



Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits

ARIES Example (Run Time)



T-Table
T1, Running, 7
T3, Running, 6

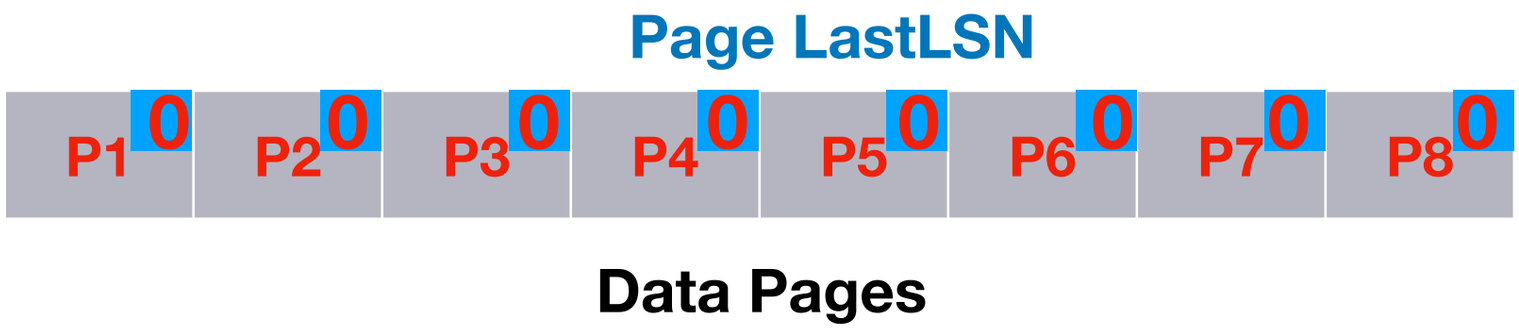
DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail
8: T1 Updates P6

FlushedLSN: 7

Main Memory (Volatile)

Hard Disk (Non-Volatile)



Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

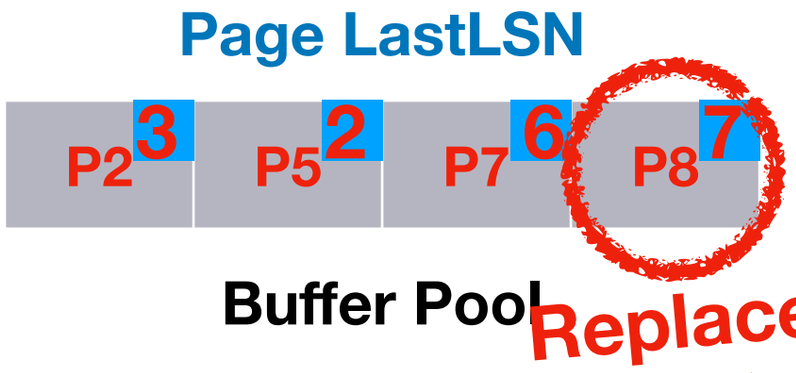
Slides by Immanuel Trummer, Cornell University

ARIES Example (Run Time)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6

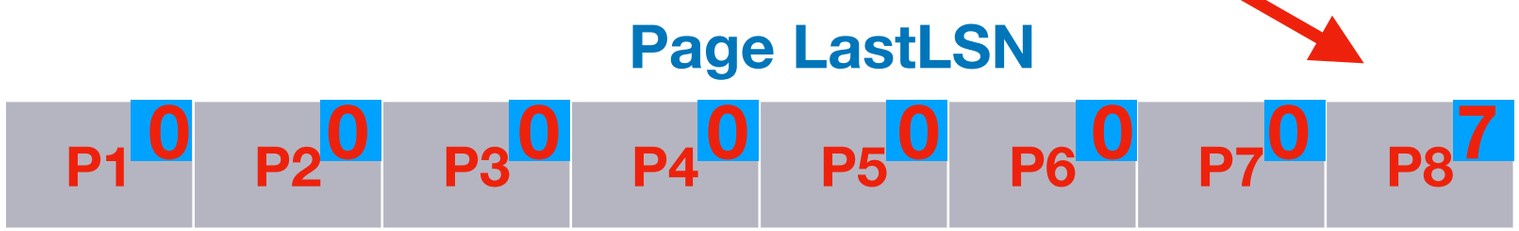
Log Tail
8: T1 Updates P6



FlushedLSN: 7

Main Memory (Volatile)

Hard Disk (Non-Volatile)



Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Run Time)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6

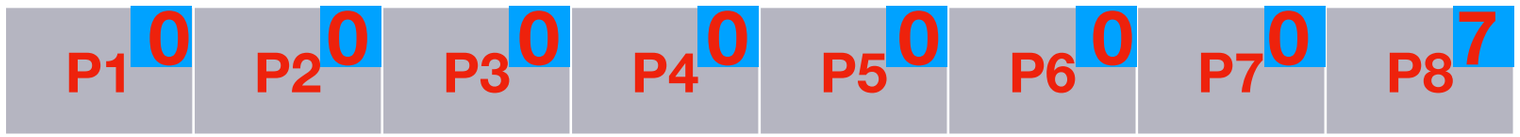
Log Tail
8: T1 Updates P6

FlushedLSN: 7

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

Outlook

- ARIES data structures
- ARIES run time behavior
- **ARIES recovery algorithm**

Recovery Phase Overview

- **Analysis phase**
  - Read log to restore transaction & dirty page tables
- **Redo phase**
  - Redo non-persisted changes of all transactions
- **Undo phase**
  - Undo changes of aborted transactions

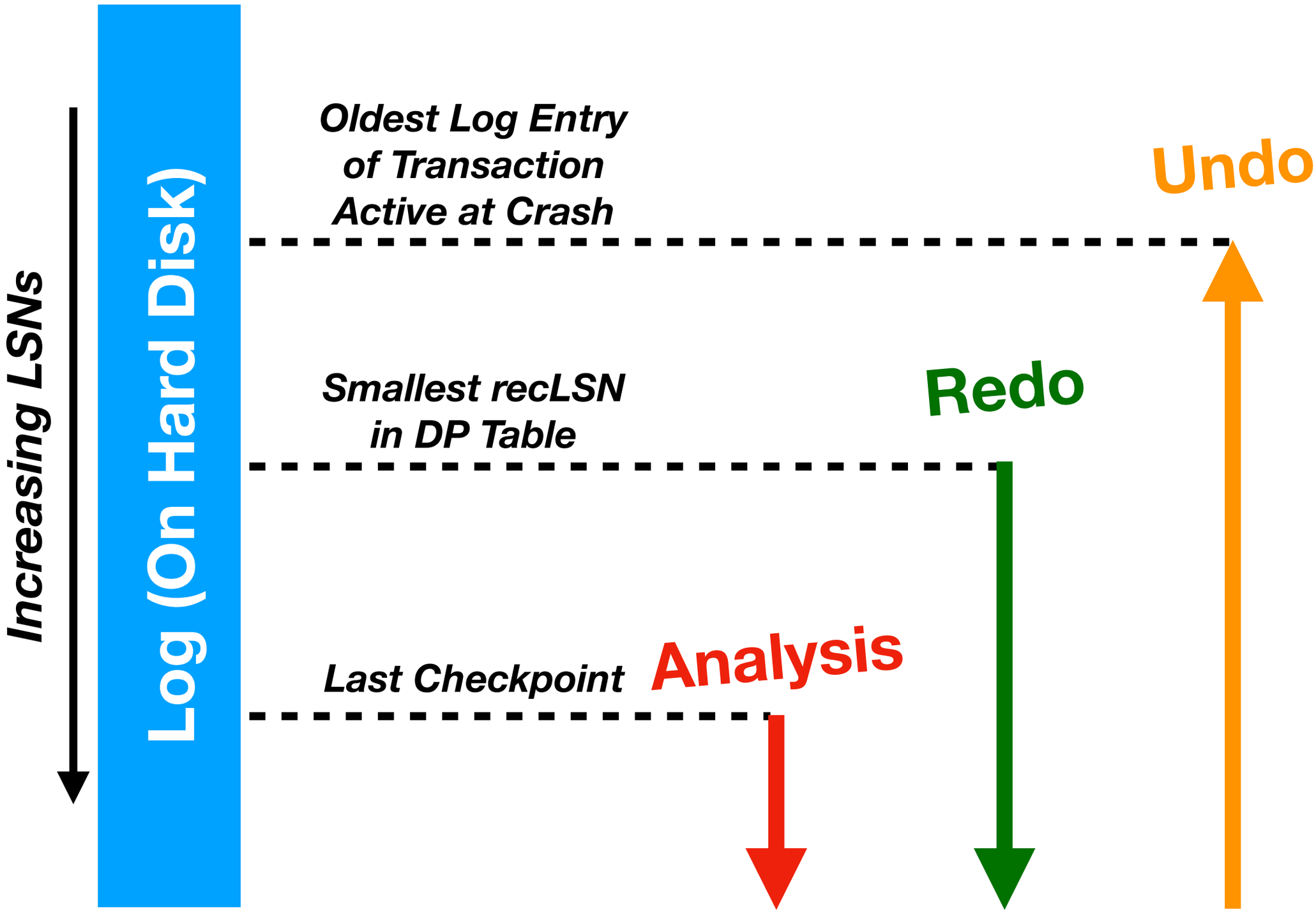
Analysis Refinement: Checkpoints

- Restore **transaction & dirty page** table during analysis
  - Without refinements, would have to read **entire log**
- **Checkpoints save time** during the analysis phase
  - Checkpoint captures **transaction & DP table** state
  - Checkpoint written to **log**
  - **Master** stores LSN of last checkpoint

Writing Checkpoints

- Want to write checkpoints **without stopping** transactions
- Log "**Begin_Checkpoint**" at checkpoint start
- Now start writing **Transaction & DP table** to log
- Log "**End_Checkpoint**" once this is done
- Logged tables represent state **at checkpoint start**

Log Scans During Recovery



Slides by Immanuel Trummer, Cornell University

Analysis Phase I

- Initialize **transaction & DP table** to last checkpoint
- Scan log forward starting from **Begin_Checkpoint**
- LSN **Lx**: Transaction **Ty** updates page **Pz**
  - If Pz not in **DP table**, add with recLSN = Lx
  - If Ty not in **transaction table**, add as running
  - Set lastLSN = Lx in **transaction table**

Analysis Phase II

- LSN Lx: transaction Ty **commits**
  - Mark Ty as **committed** in transaction table
- LSN Lx: transaction Ty **aborts**
  - Mark Ty as **aborted** in transaction table
- LSN Lx: **end record** for transaction Ty
  - **Remove** Ty from transaction table

Do We Get Precise Transaction Table & DP Table State Before Crash?

ARIES Example (Analysis)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6

Log Tail
8: T1 Updates P6

Page LastLSN

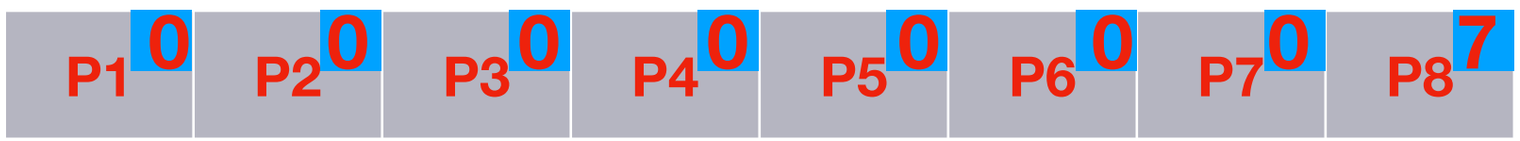
Buffer Pool

FlushedLSN: 7

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Analysis)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6

Log Tail
8: T1 Updates P6

Page LastLSN

Buffer Pool

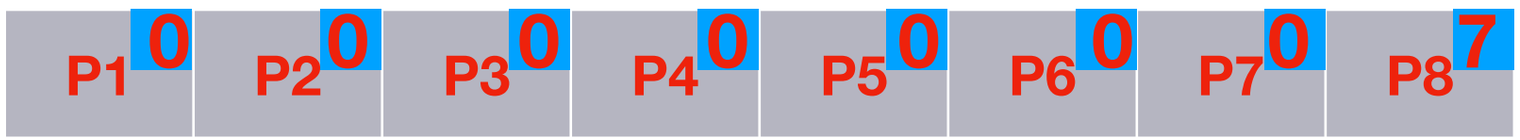
FlushedLSN: 7

Crash!

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



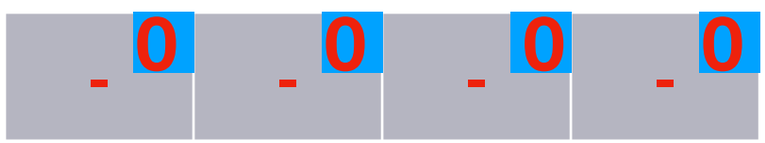
Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Analysis)

T-Table	DP Table	Log Tail

Page LastLSN



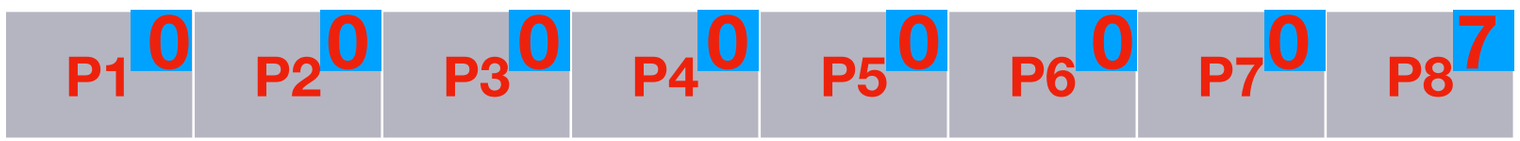
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



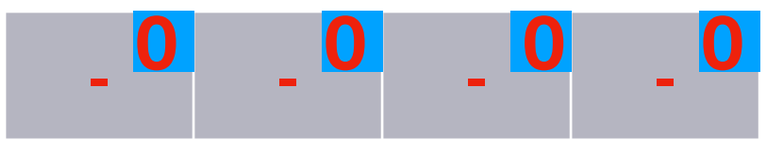
Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Analysis)

T-Table	DP Table	Log Tail
T1, Running, 1	P2, 1	

Page LastLSN



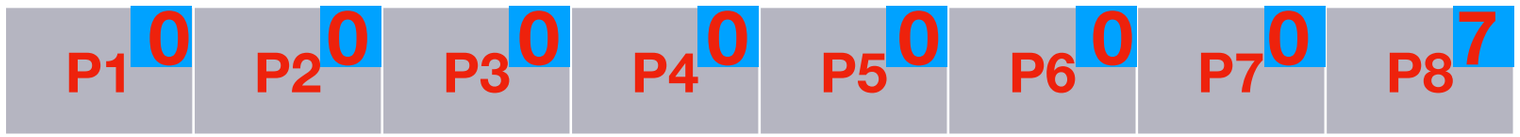
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

Slides by Immanuel Trummer, Cornell University

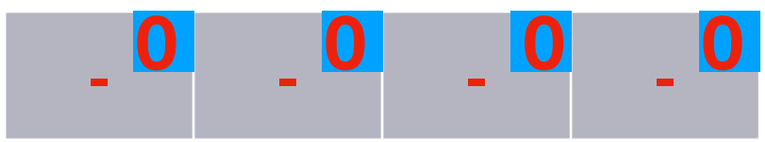
ARIES Example (Analysis)

T-Table
T1, Running, 1
T2, Running, 2

DP Table
P2, 1
P5, 2

Log Tail

Page LastLSN



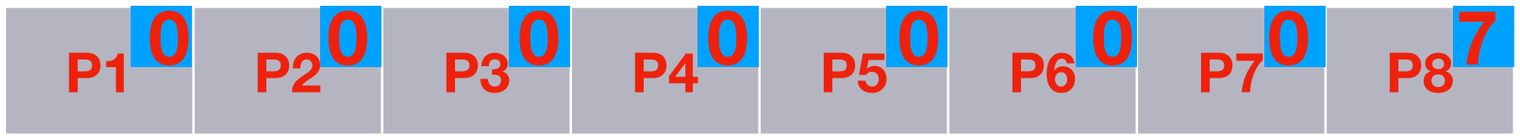
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

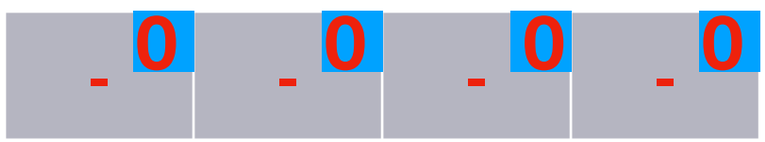
ARIES Example (Analysis)

T-Table
T1, Running, 3
T2, Running, 2

DP Table
P2, 1
P5, 2

Log Tail

Page LastLSN



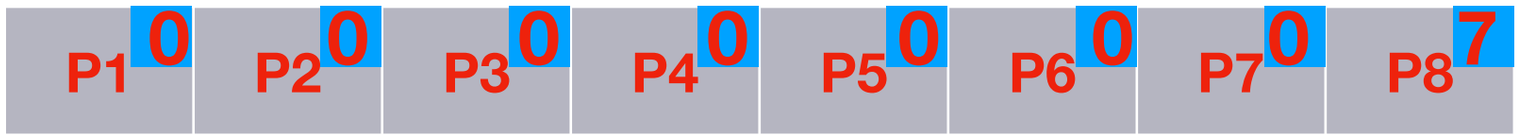
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

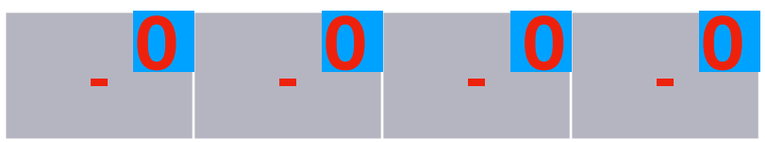
ARIES Example (Analysis)

T-Table
T1, Running, 3
T2, Committed, 4

DP Table
P2, 1
P5, 2

Log Tail

Page LastLSN



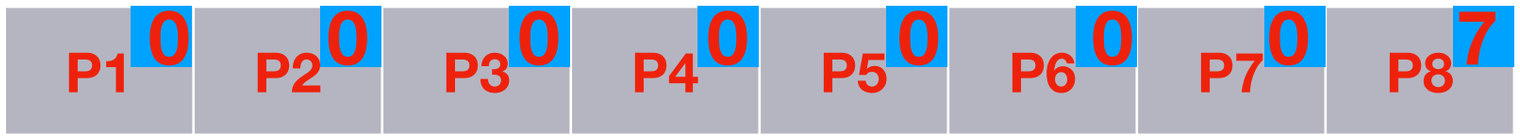
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



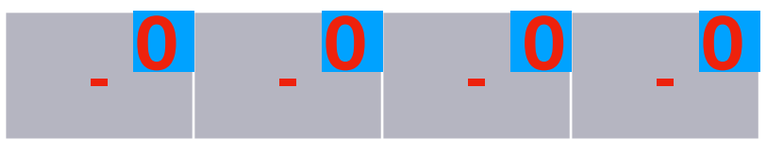
Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Analysis)

T-Table	DP Table	Log Tail
T1, Running, 3	P2, 1	
	P5, 2	

Page LastLSN



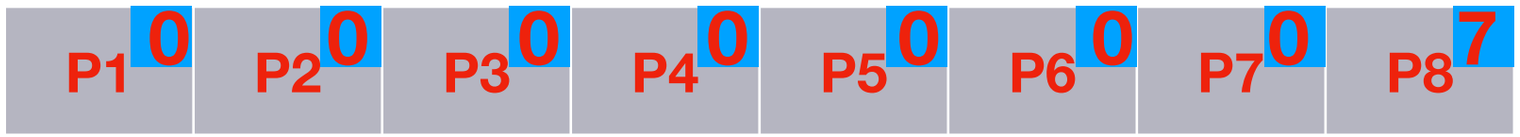
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



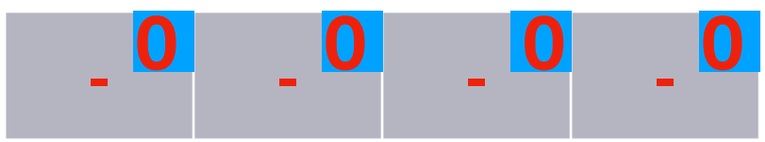
Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

Slides by Immanuel Trummer, Cornell University

ARIES Example (Analysis)

Page LastLSN



Buffer Pool

T-Table
T1, Running, 3
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6

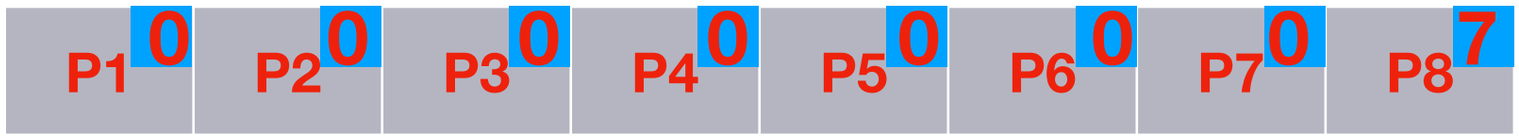
Log Tail

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

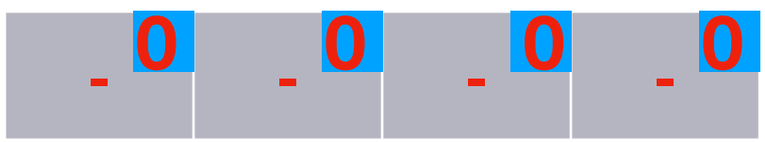


Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Analysis)

Page LastLSN



Buffer Pool

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

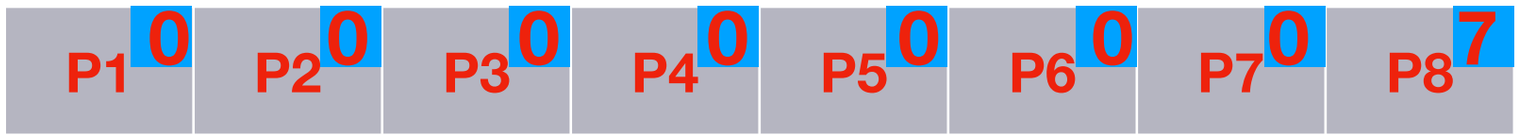
Log Tail

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

Redo Phase

- Start scanning log from **earliest recLSN** in DP table
  - This is **first log entry** that made some page dirty
- Focus on **data updates** and **compensation log records**
  - **Redo change**, update pageLSN (no new log entry)
  - May **avoid redo** if certain conditions hold (see next)

Avoiding Redo Operations

- We can avoid a redo if one of those **conditions** holds

1. Page affected by update is **not in dirty page table**
2. Affected page in DP table but with **recLSN > LSN**
3. PageLSN of page on hard disk with **pageLSN ≥ LSN**

Why Do We Check Conditions in This Order?

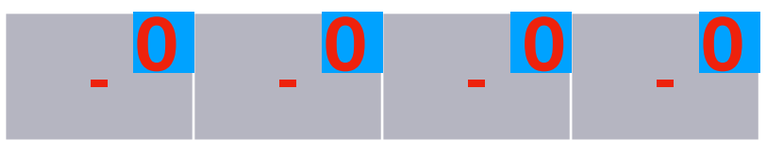
ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN



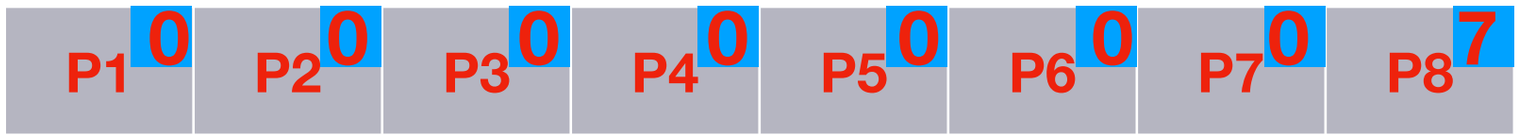
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Redo)

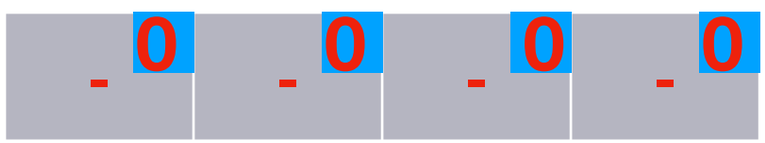
T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Earliest RecLSN

Page LastLSN



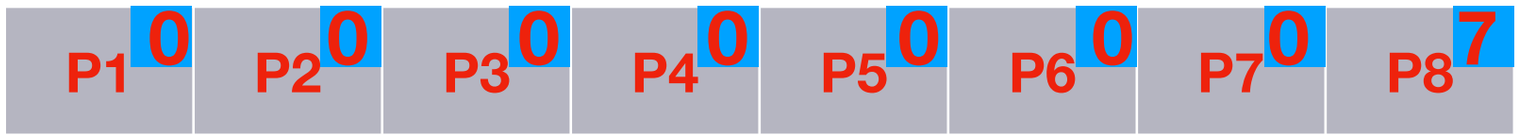
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Redo)

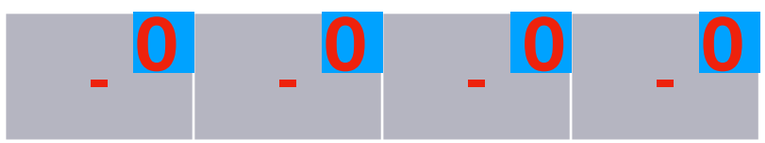
T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Earliest RecLSN

Page LastLSN



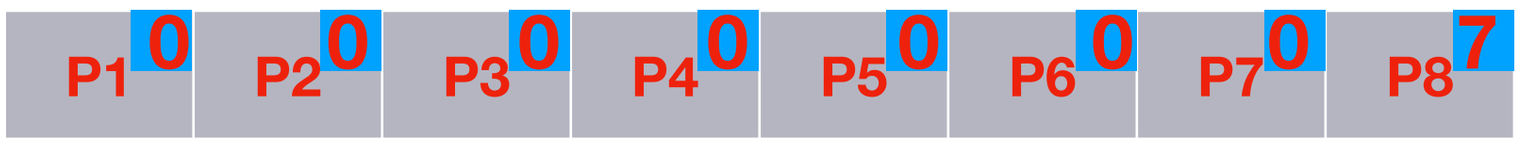
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

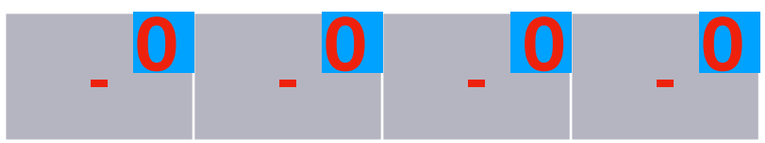
ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN



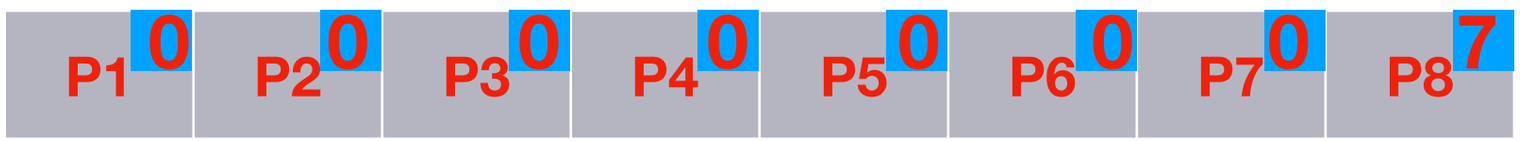
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

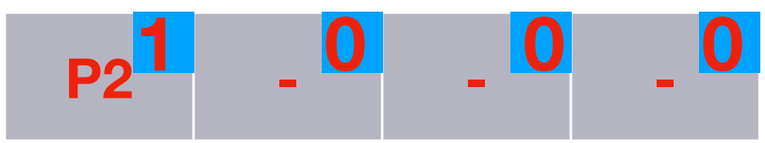
ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN



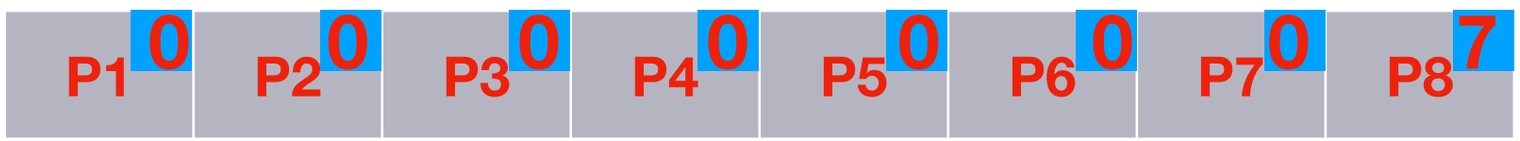
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Redo)

Page LastLSN

P2 1	P5 2	- 0	- 0

Buffer Pool

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN

P1 0	P2 0	P3 0	P4 0	P5 0	P6 0	P7 0	P8 7

Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN

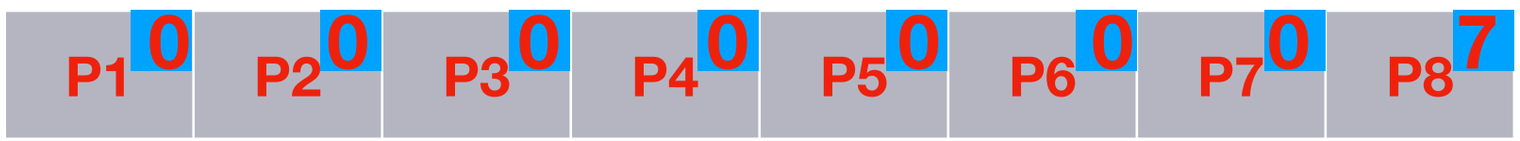
Buffer Pool

FlushedLSN: 0

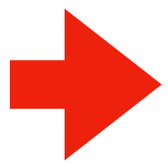
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN

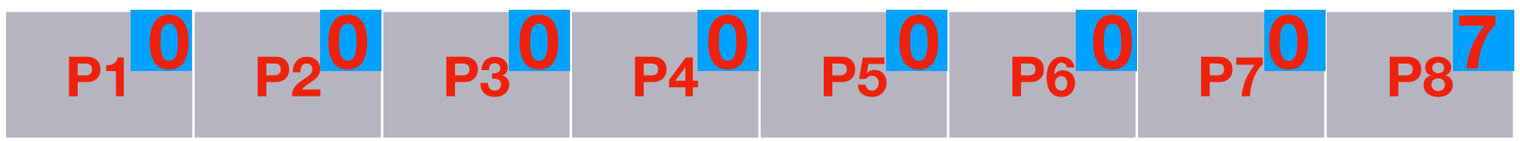
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN

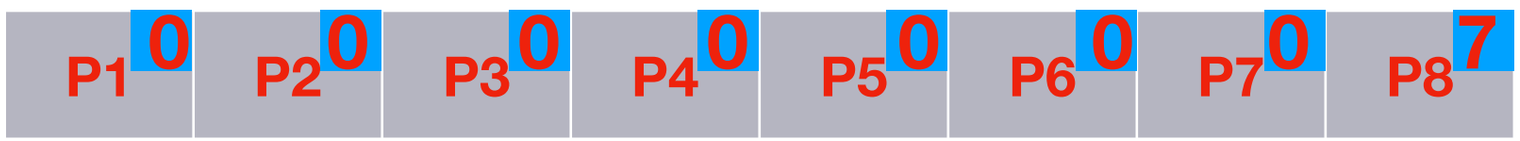
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN

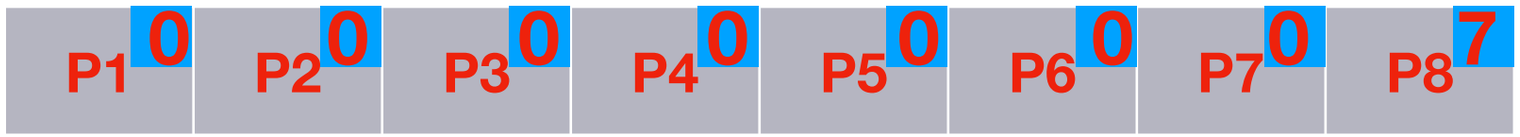
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

Slides by Immanuel Trummer, Cornell University

ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN

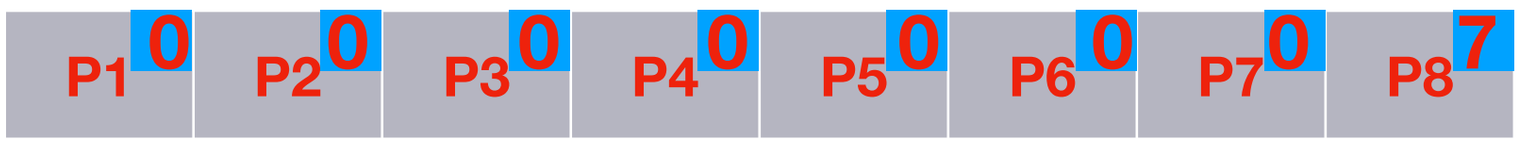
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

ARIES Example (Redo)

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail

Page LastLSN

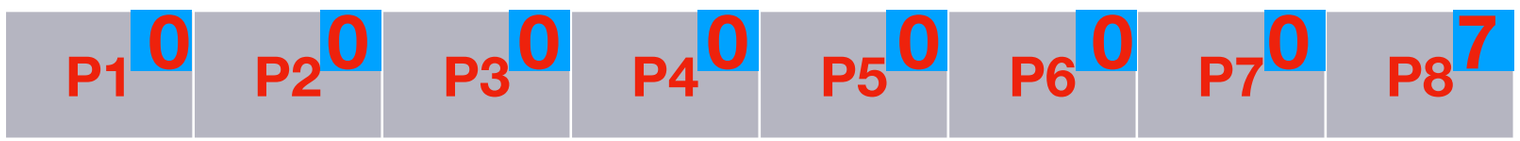
Buffer Pool

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

Do We Need To Redo This ... ?

Slides by Immanuel Trummer, Cornell University

Undo Phase

- Undo changes by transactions **running at crash**
- Undo changes in **reverse log order**
  - Follow **prevLSN pointers** backwards in log
  - Write **compensation log records** while undoing

ARIES Example (Undo)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

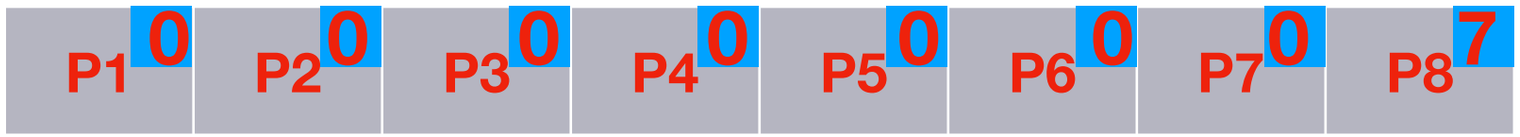
Log Tail

FlushedLSN: 0

Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages

Log
1: T1 Updates P2
2: T2 Updates P5
3: T1 Updates P2
4: T2 Commits
5: T2 End
6: T3 Updates P7
7: T1 Updates P8

Slides by Immanuel Trummer, Cornell University

ARIES Example (Undo)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 7
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

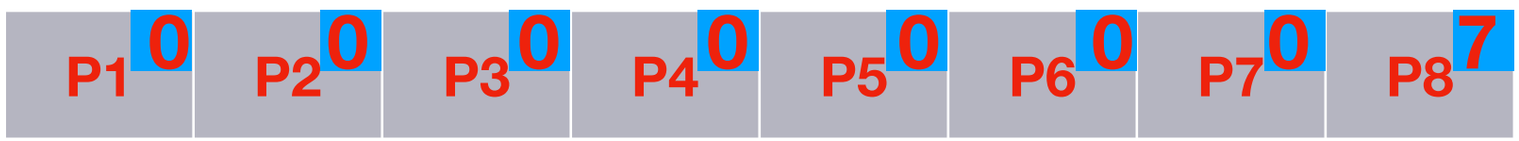
Log Tail

FlushedLSN: 0

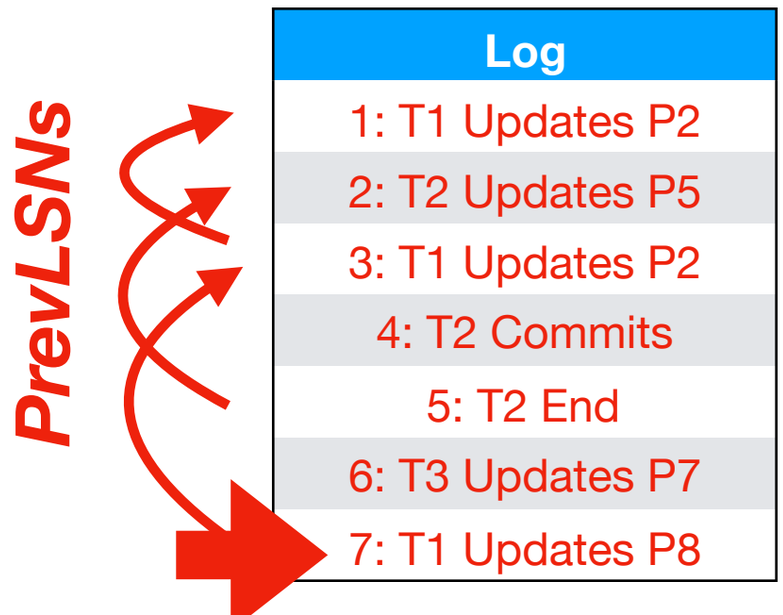
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



ARIES Example (Undo)

T-Table
T1, Running, 8
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

Log Tail
8: CLR T1 P8

Page LastLSN

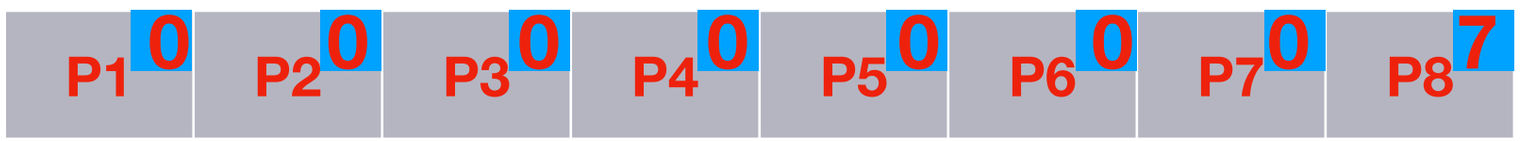
Buffer Pool

FlushedLSN: 0

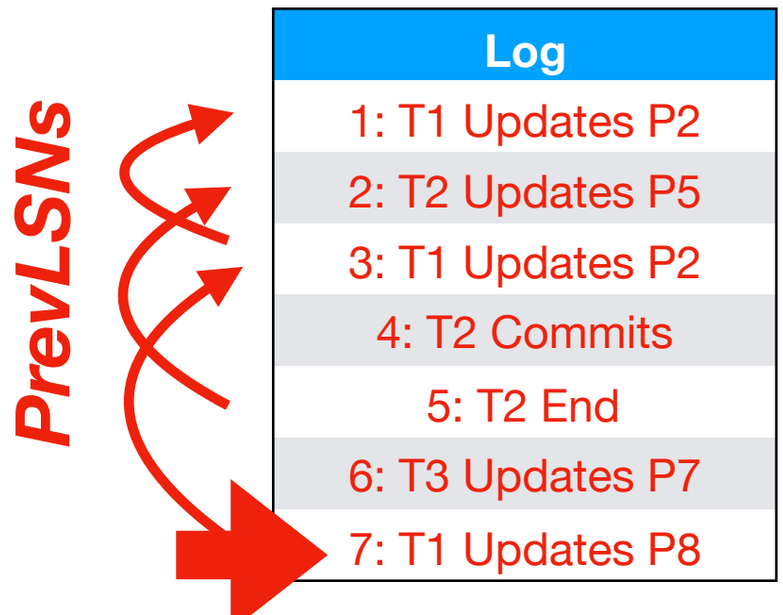
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



ARIES Example (Undo)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 8
T3, Running, 6

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

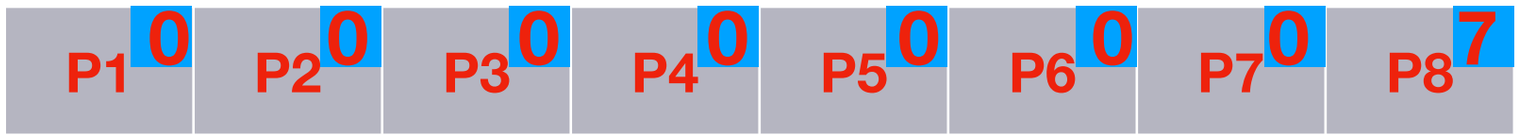
Log Tail
8: CLR T1 P8

FlushedLSN: 0

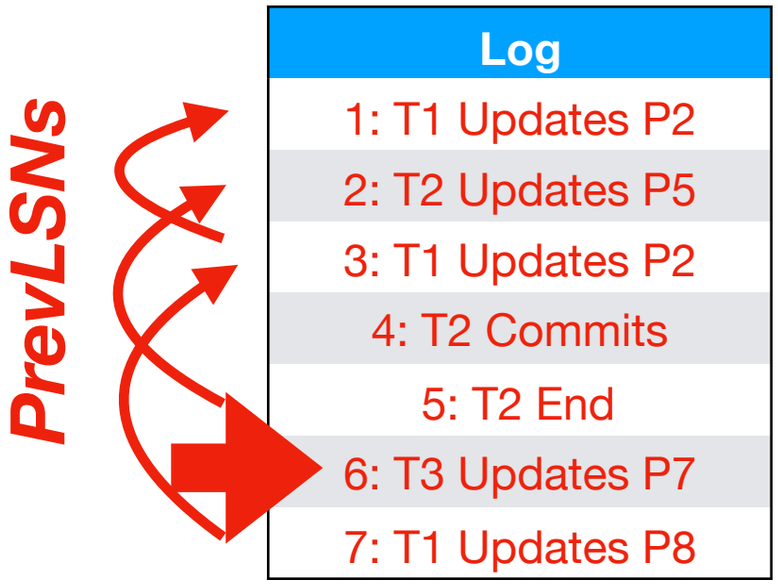
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



ARIES Example (Undo)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 8
T3, Running, 9

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

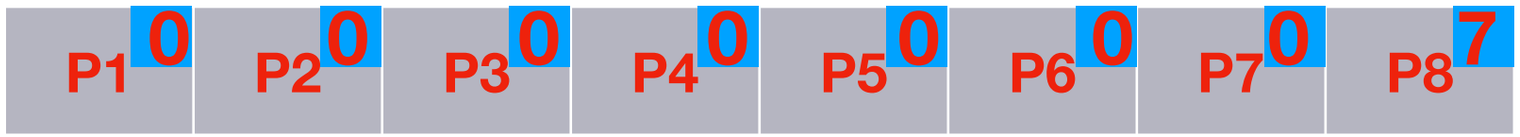
Log Tail
8: CLR T1 P8
9: CLR T3 P7

FlushedLSN: 0

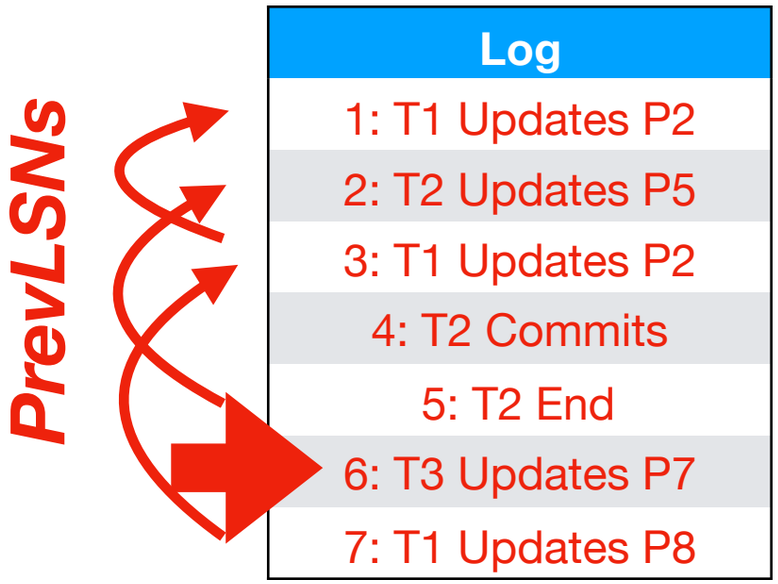
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



ARIES Example (Undo)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 8
T3, Running, 9

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

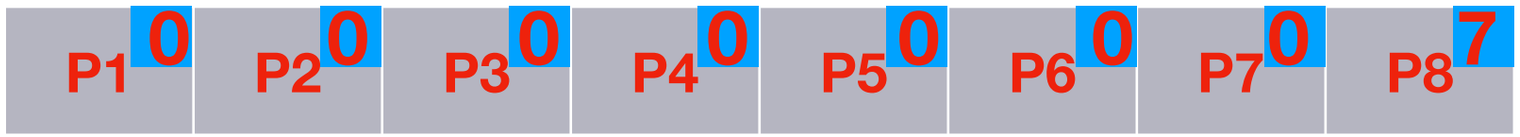
Log Tail
8: CLR T1 P8
9: CLR T3 P7

FlushedLSN: 0

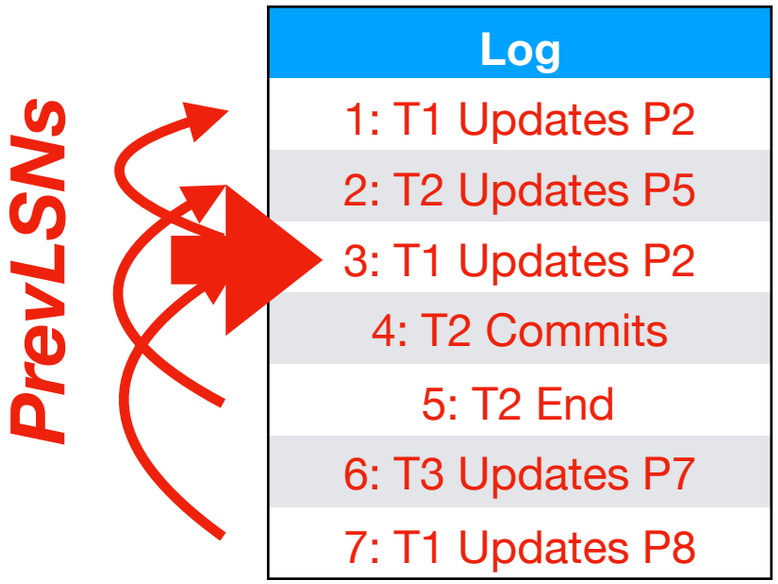
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



ARIES Example (Undo)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 10
T3, Running, 9

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

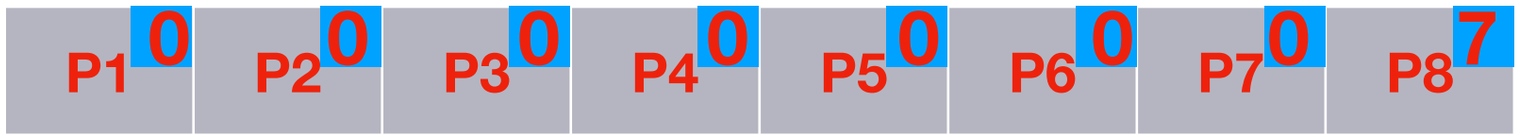
Log Tail
8: CLR T1 P8
9: CLR T3 P7
10: CLR T1 P2

FlushedLSN: 0

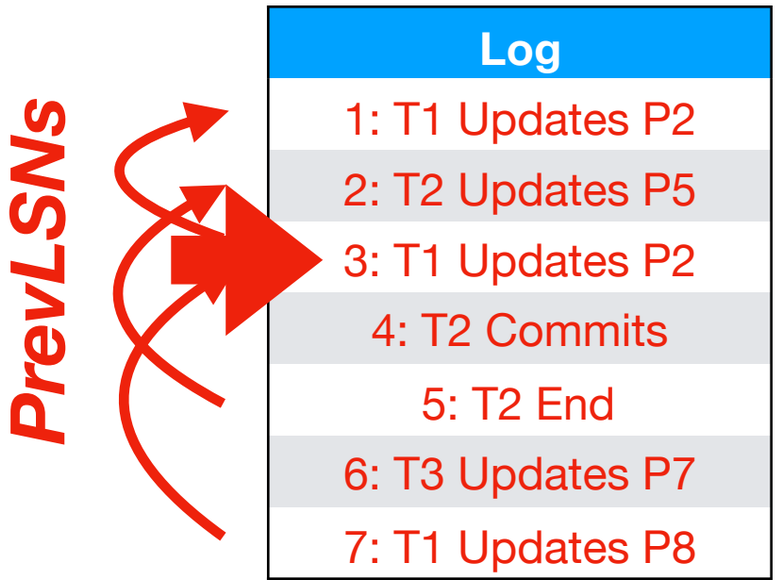
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



ARIES Example (Undo)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 10
T3, Running, 9

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

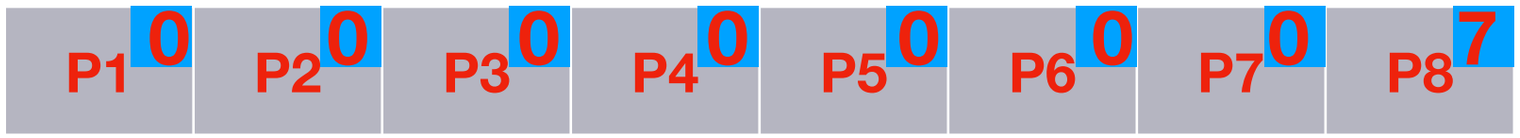
Log Tail
8: CLR T1 P8
9: CLR T3 P7
10: CLR T1 P2

FlushedLSN: 0

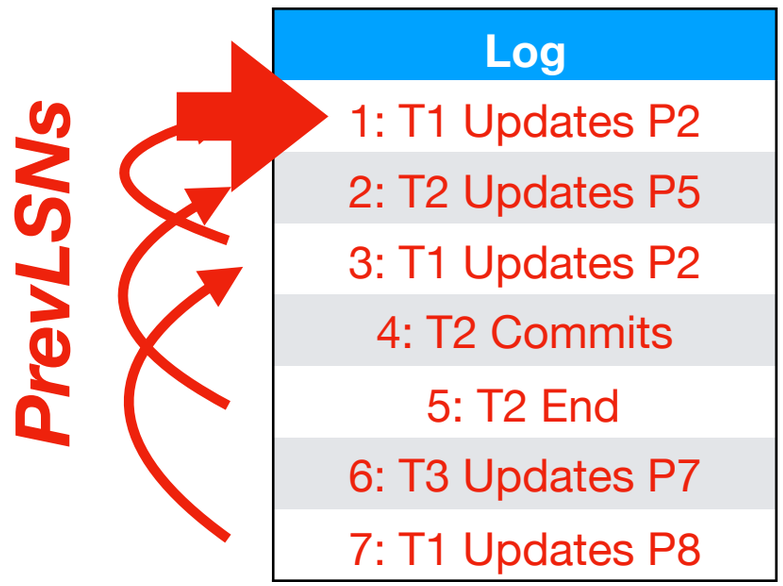
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



ARIES Example (Undo)

Page LastLSN

Buffer Pool

T-Table
T1, Running, 10
T3, Running, 9

DP Table
P2, 1
P5, 2
P7, 6
P8, 7

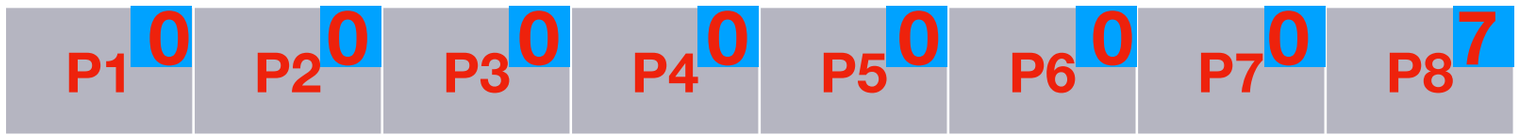
Log Tail
8: CLR T1 P8
9: CLR T3 P7
10: CLR T1 P2
11: CLR T1 P2

FlushedLSN: 0

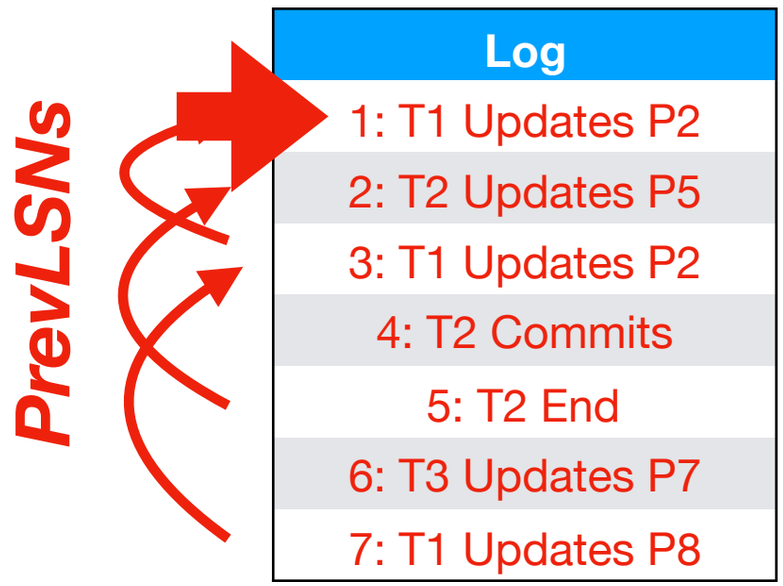
Main Memory (Volatile)

Hard Disk (Non-Volatile)

Page LastLSN



Data Pages



Slides by Immanuel Trummer, Cornell University

Crash During Recovery?

- ARIES can deal with failures during **any** recovery phase
- Crash during analysis? Simply **restart analysis**.
- Crash during redo? Restart **analysis and redo**.
  - May not have to redo everything if **changes persisted**
- Crash during undo? **Restart all phases**.
  - Note: **CLRs** are redone but never undone!