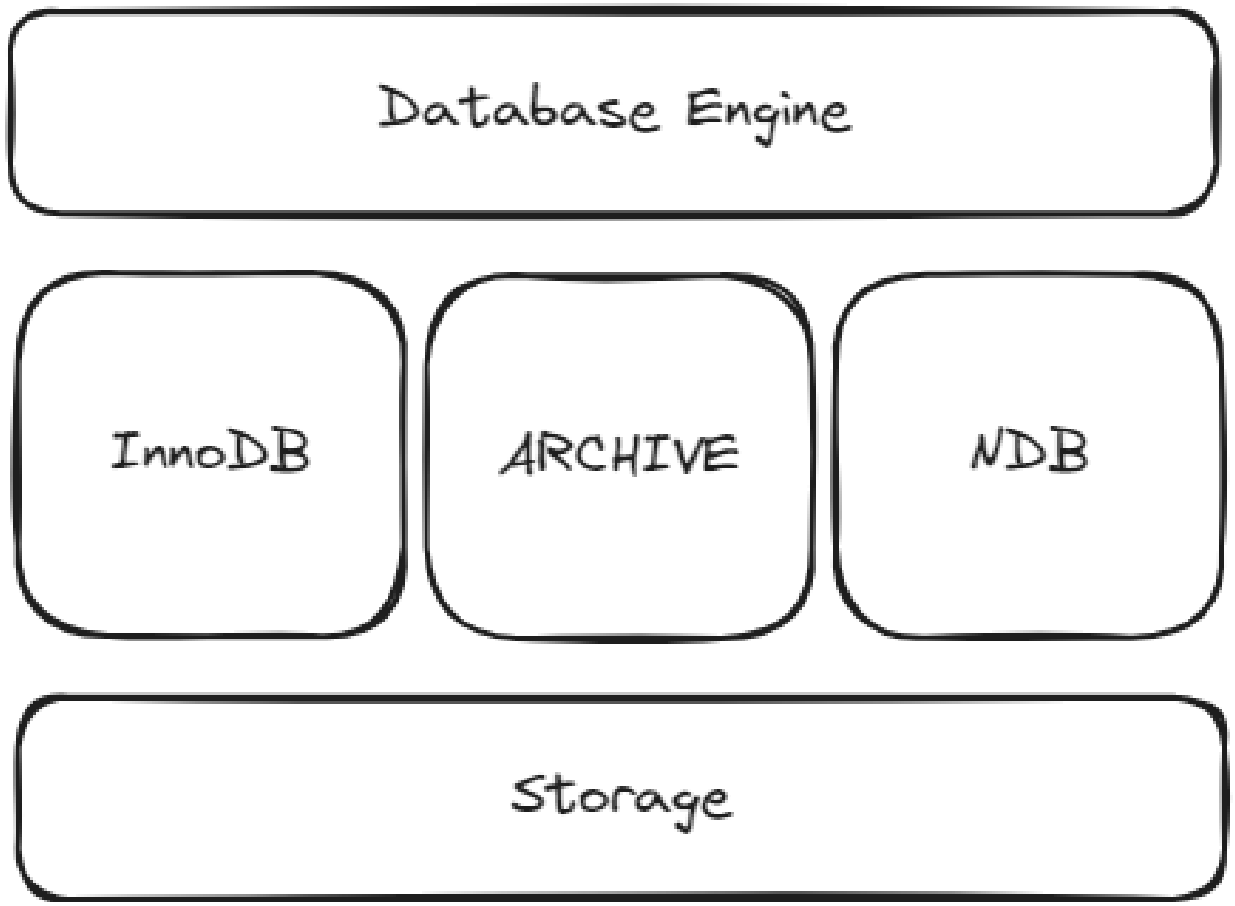
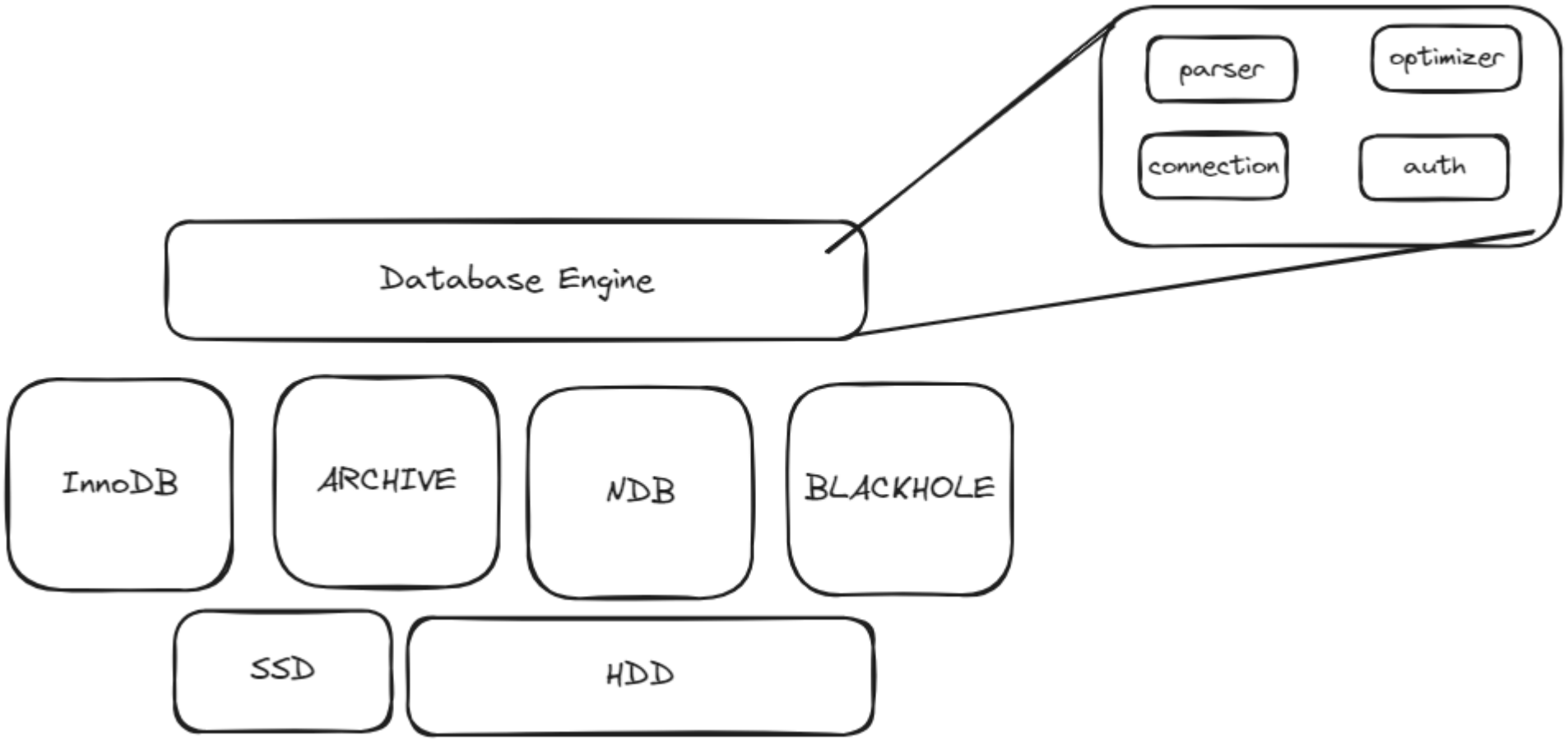
Database internals

MVCC

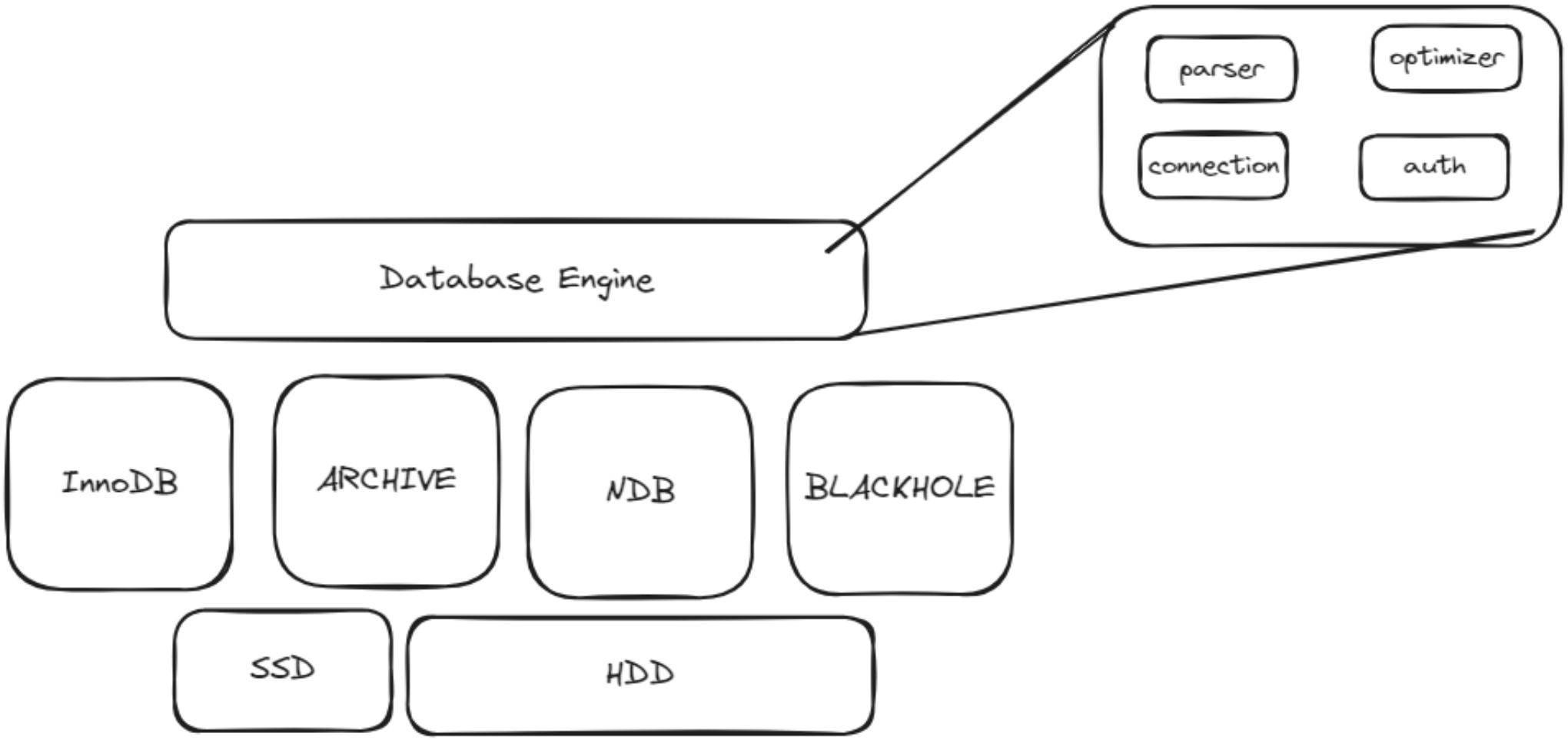
Database Engine (MySQL InnoDB)



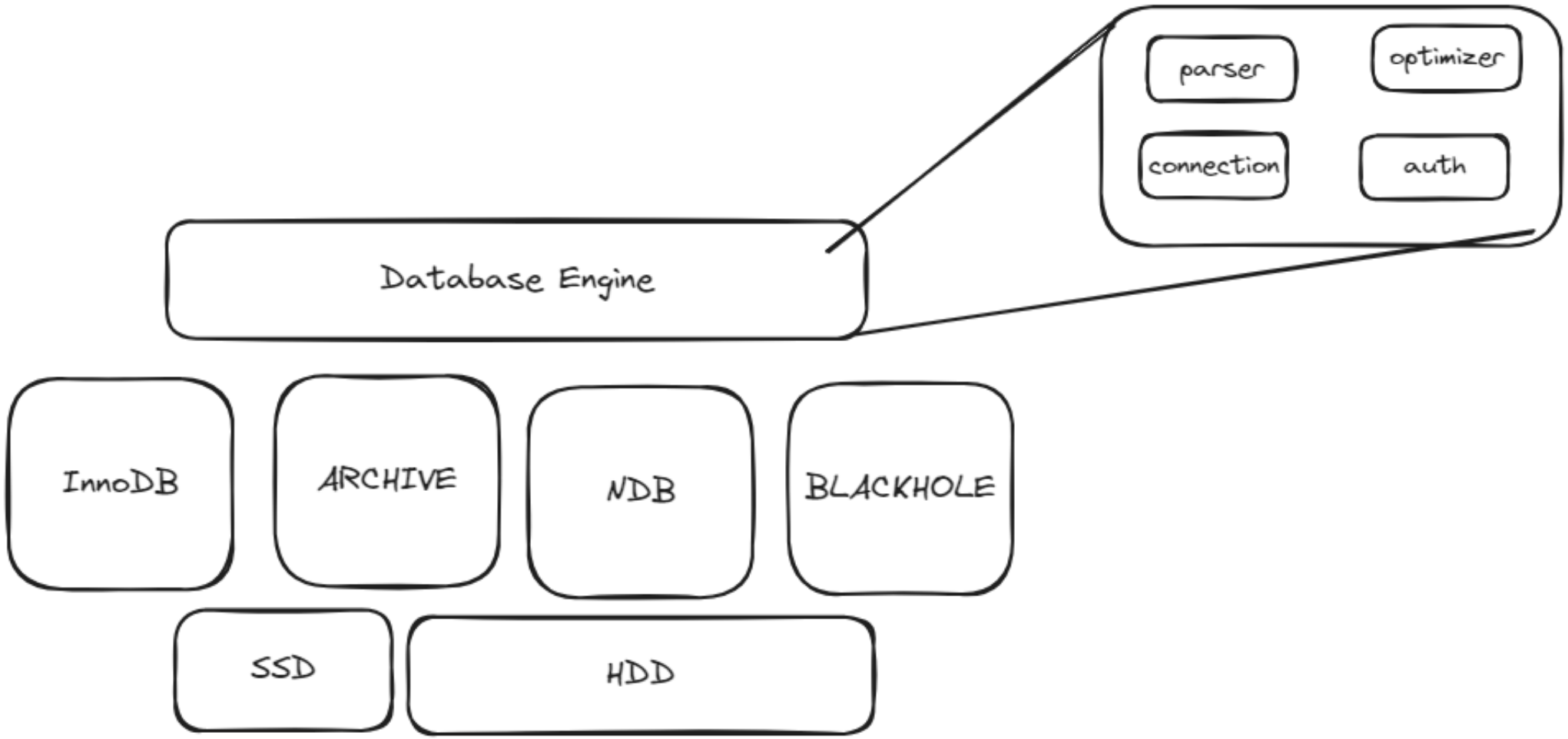
Why should we use multi storage engine?



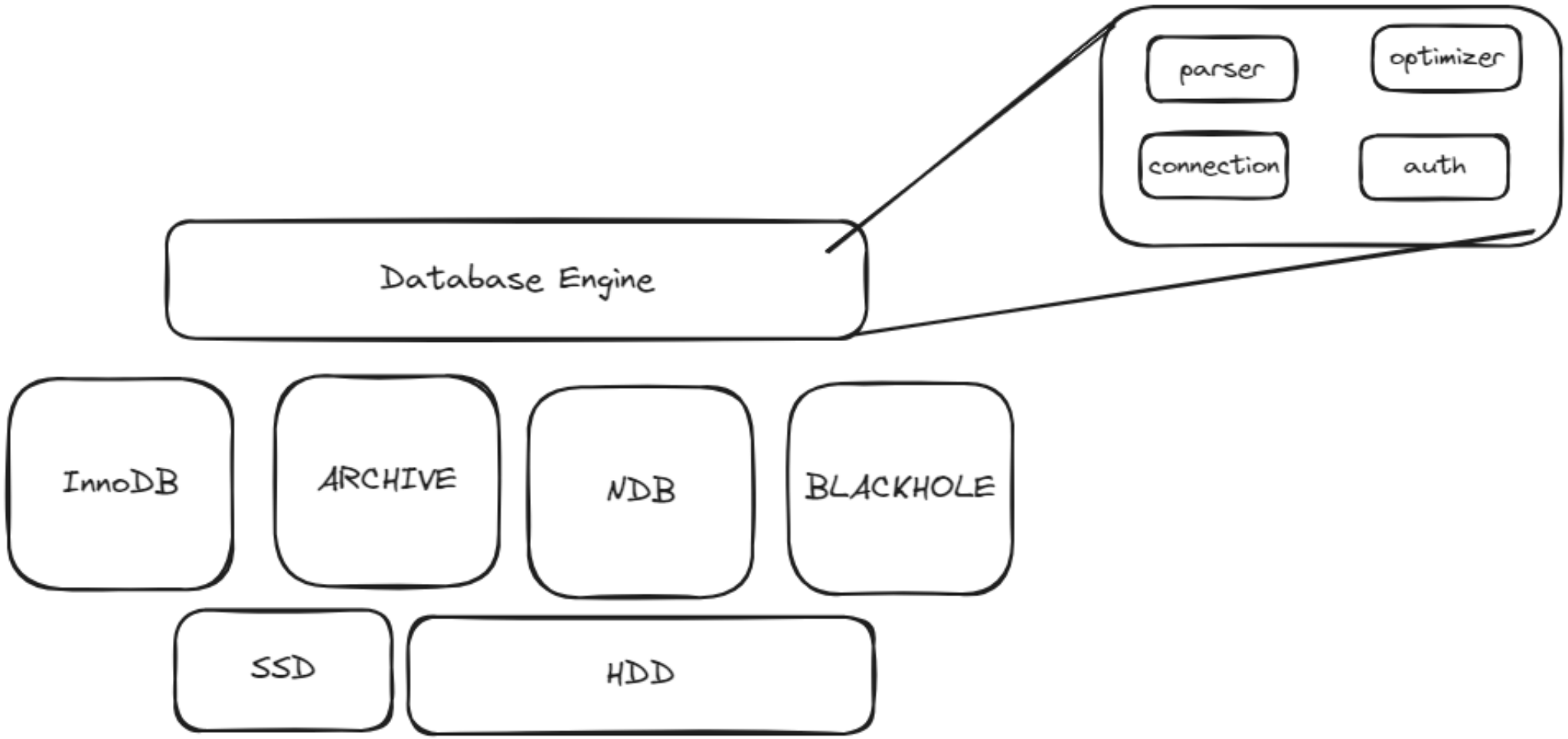
Does storage engine share storage?



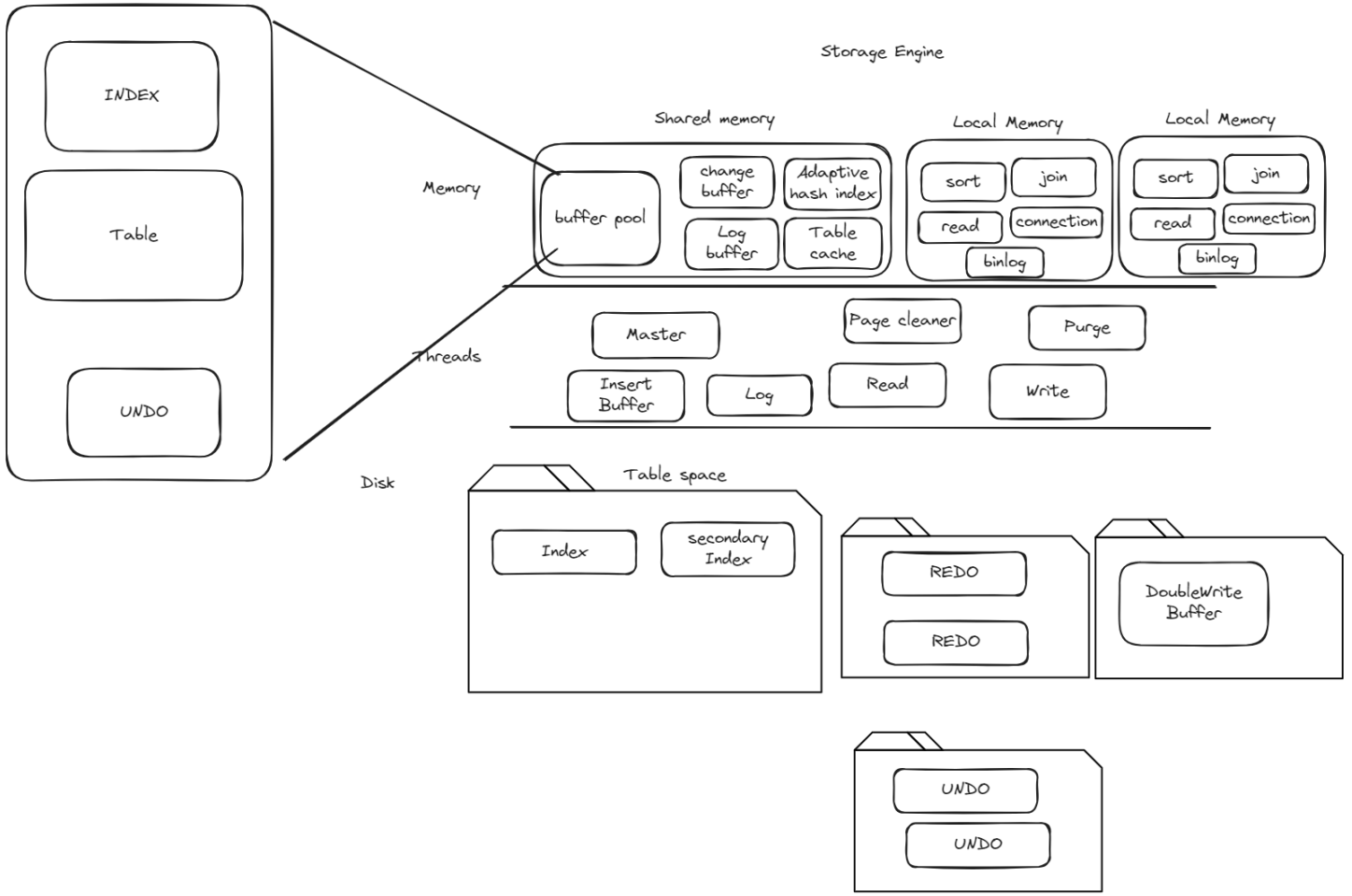
OLTP vs OLAP


parser
optimizer
connection
auth
Database Engine
InnoDB
ARCHIVE
NDB
BLACKHOLE
SSD
HDD

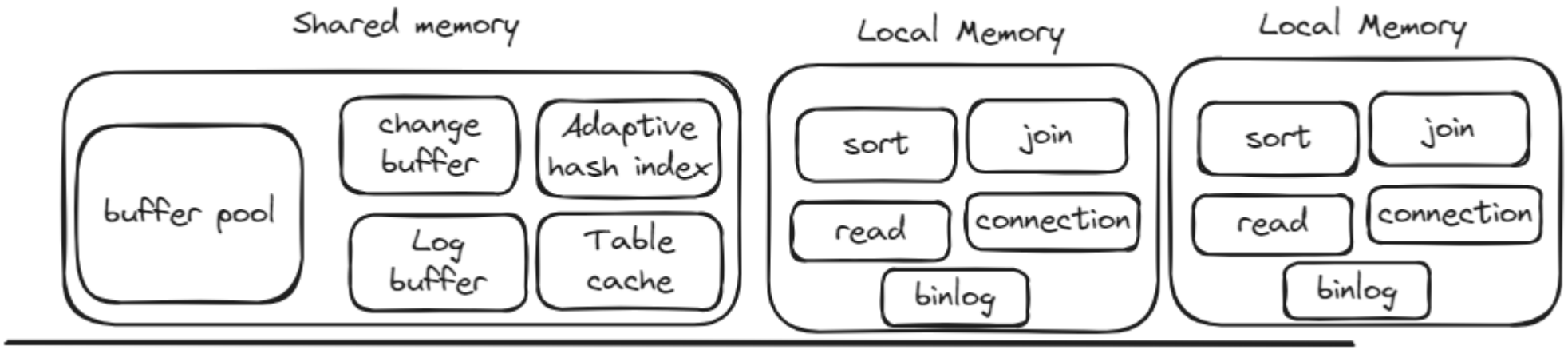
How about using memory engine instead of Redis?



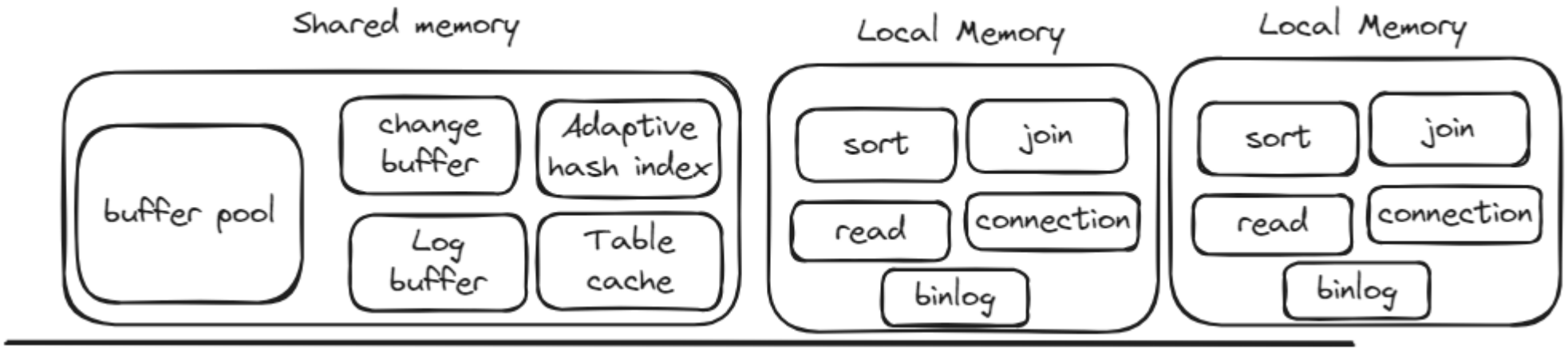
How storage engine works?



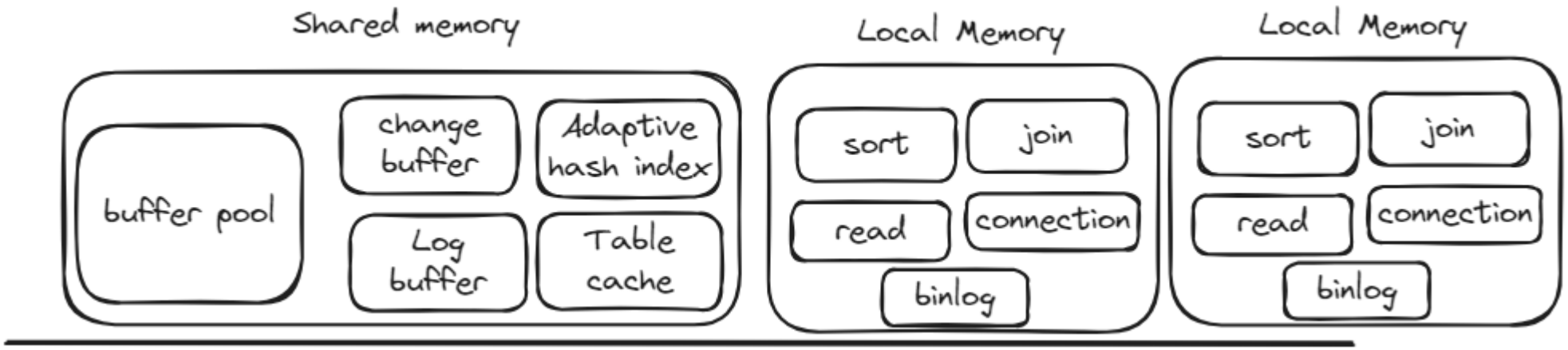
Memory


Shared memory
buffer pool
change buffer
Adaptive hash index
Log buffer
Table cache
Local Memory
sort
join
read
connection
binlog
Local Memory
sort
join
read
connection
binlog

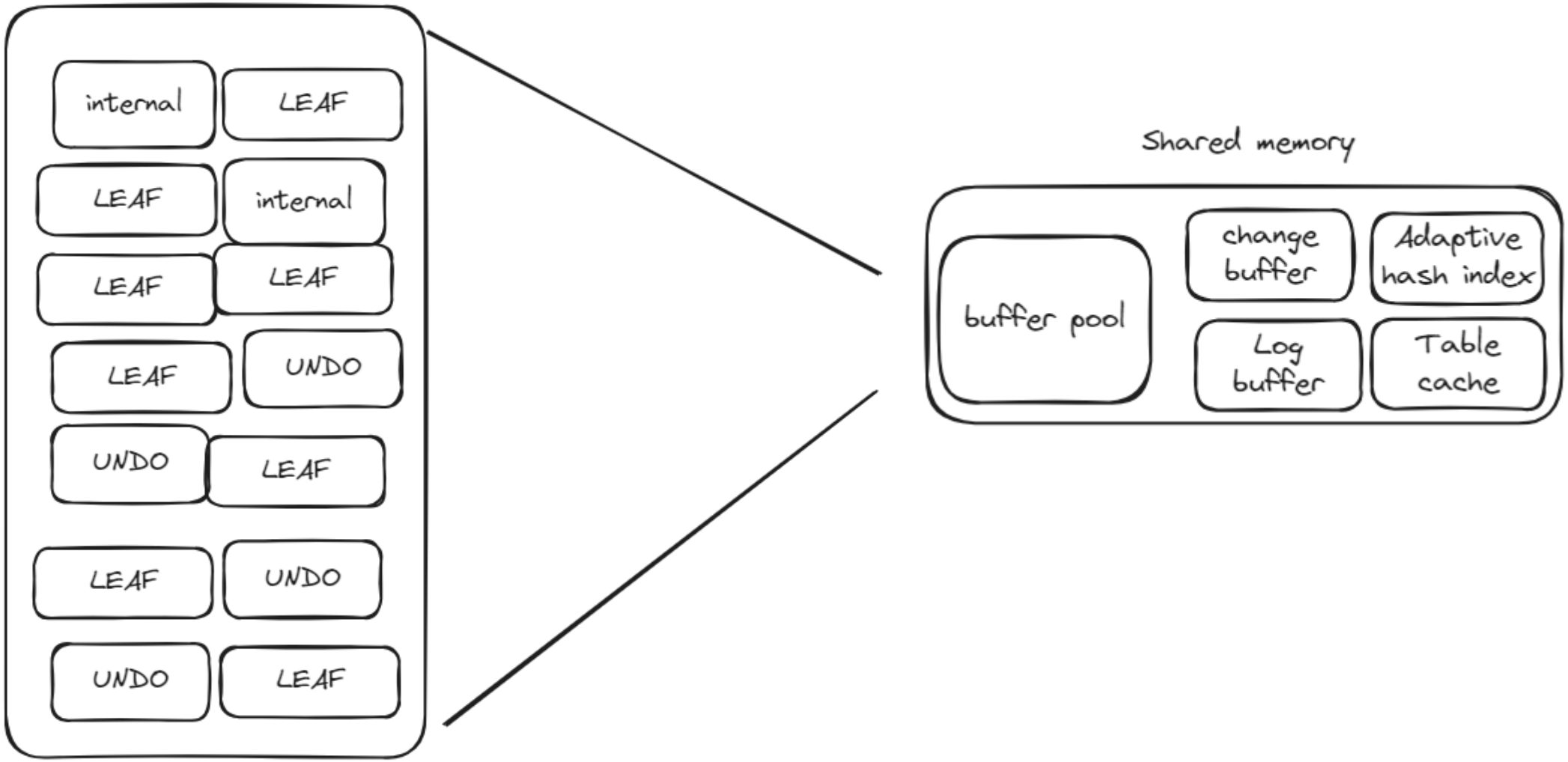
Database connection?



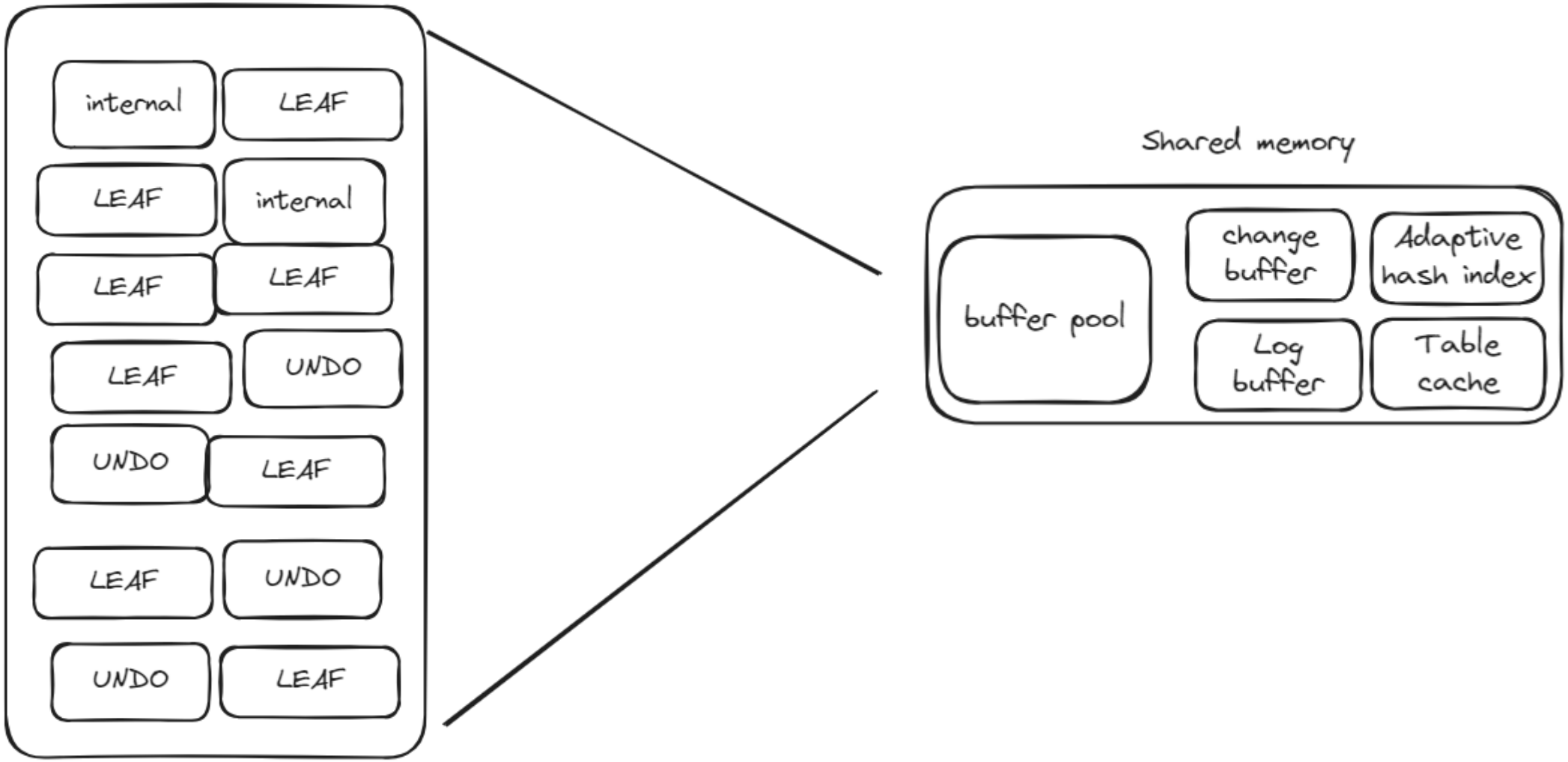
What's doing DBA?



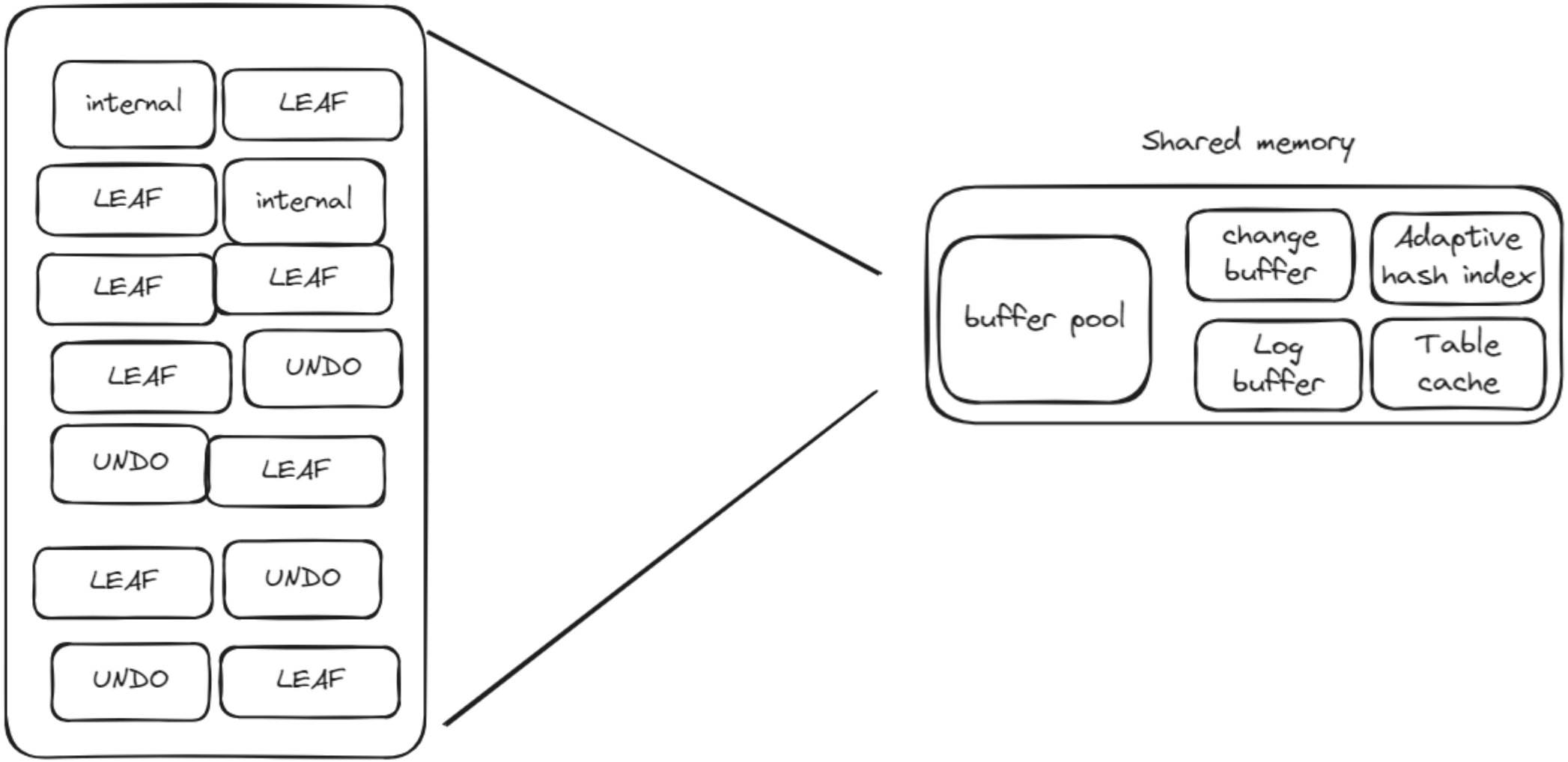
Buffer pool


internal
LEAF
LEAF
internal
LEAF
LEAF
LEAF
UNDO
UNDO
LEAF
LEAF
UNDO
UNDO
LEAF
Shared memory
buffer pool
change buffer
Adaptive hash index
Log buffer
Table cache

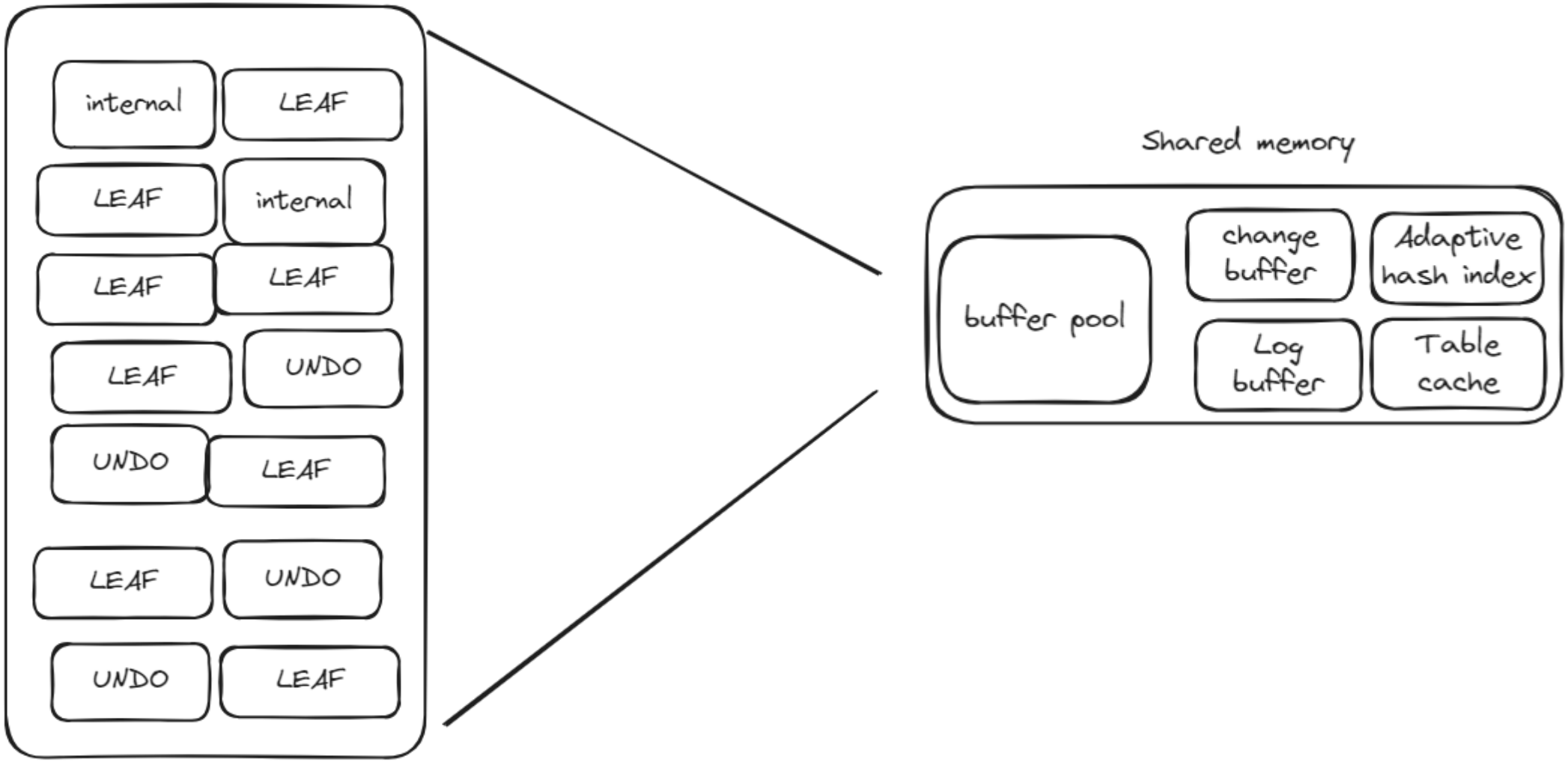
If Commit, data is written on disk?



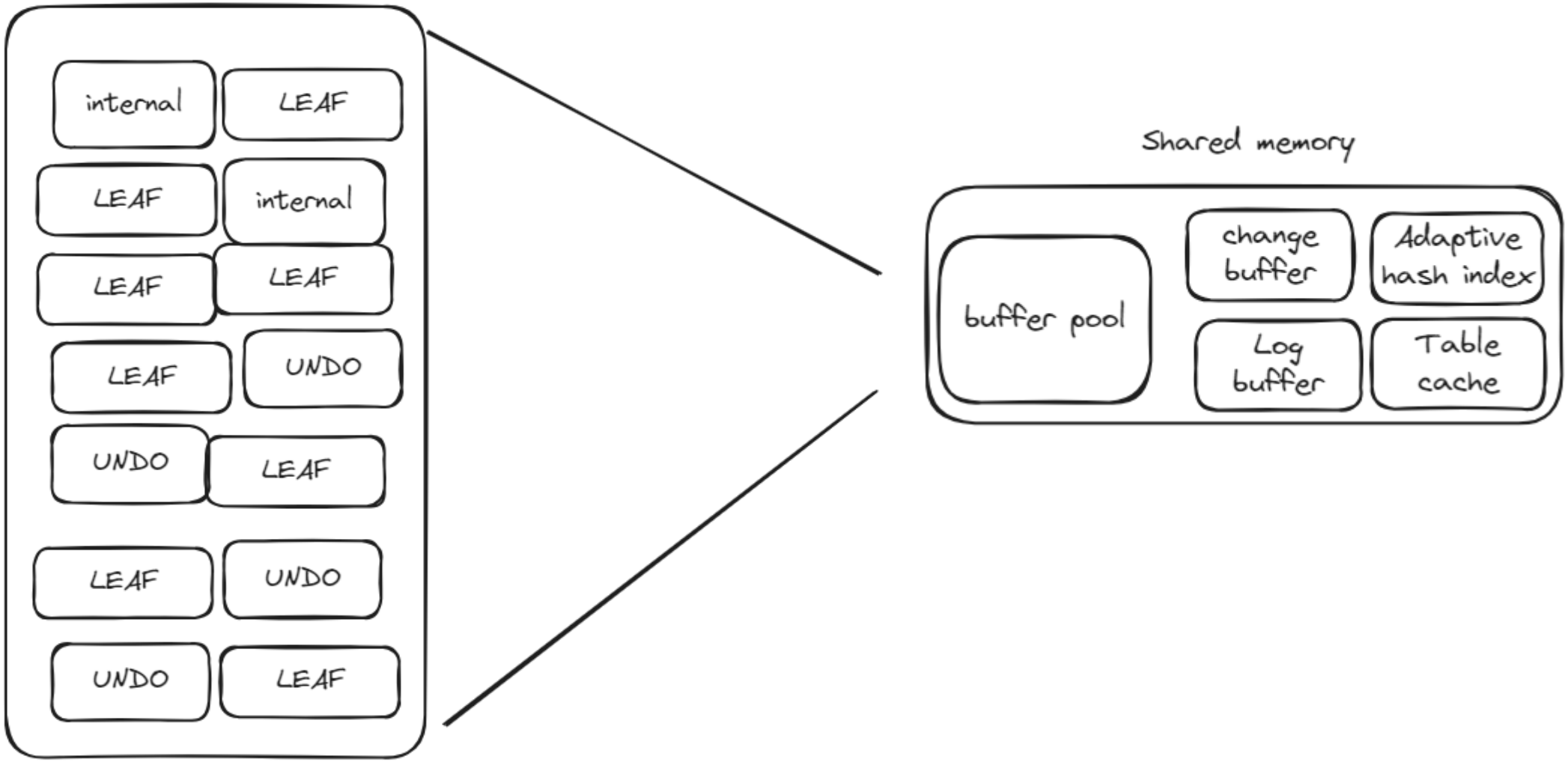
What is dirty page??



How data cleaned?



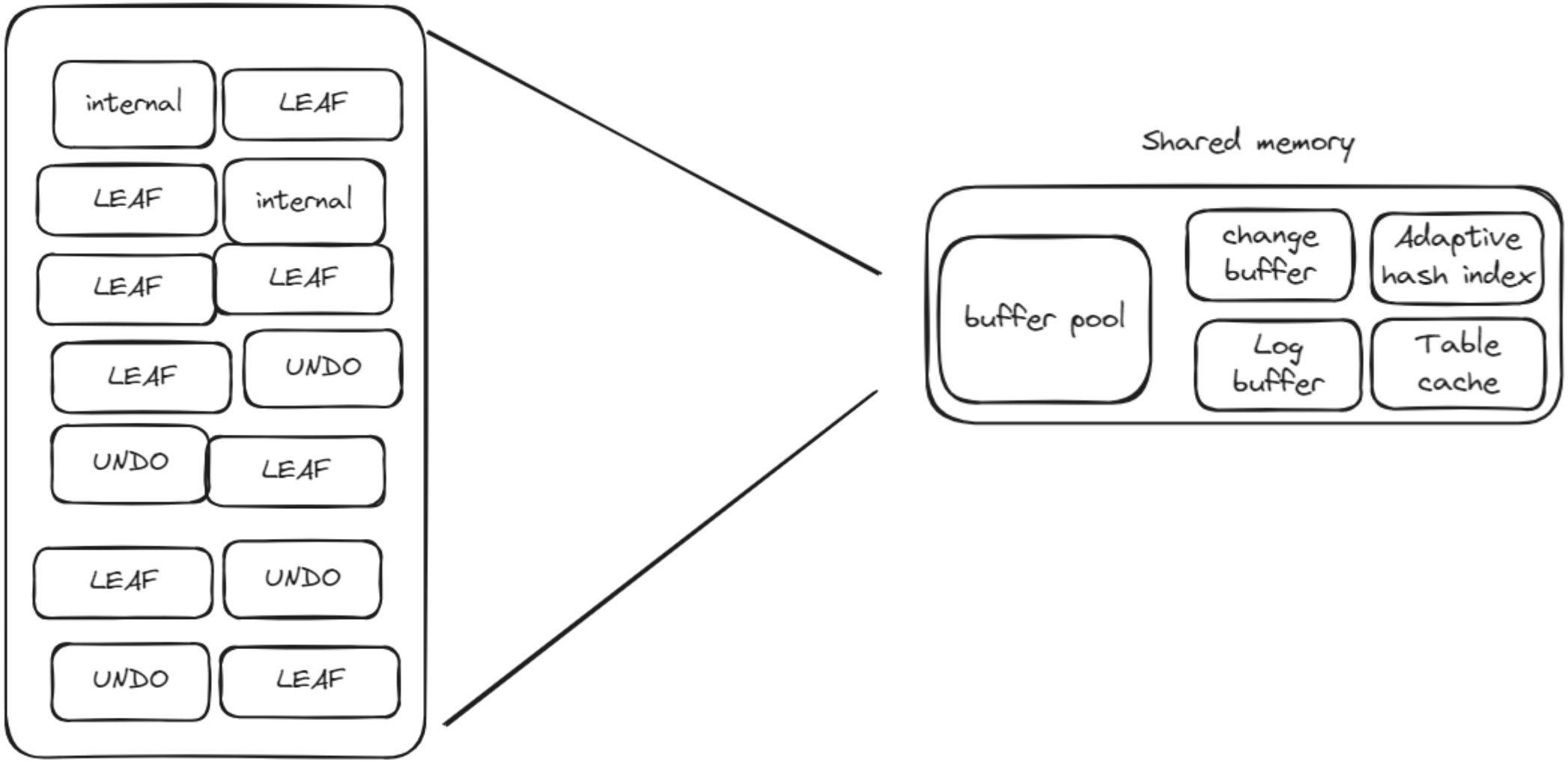
What is Undo log?



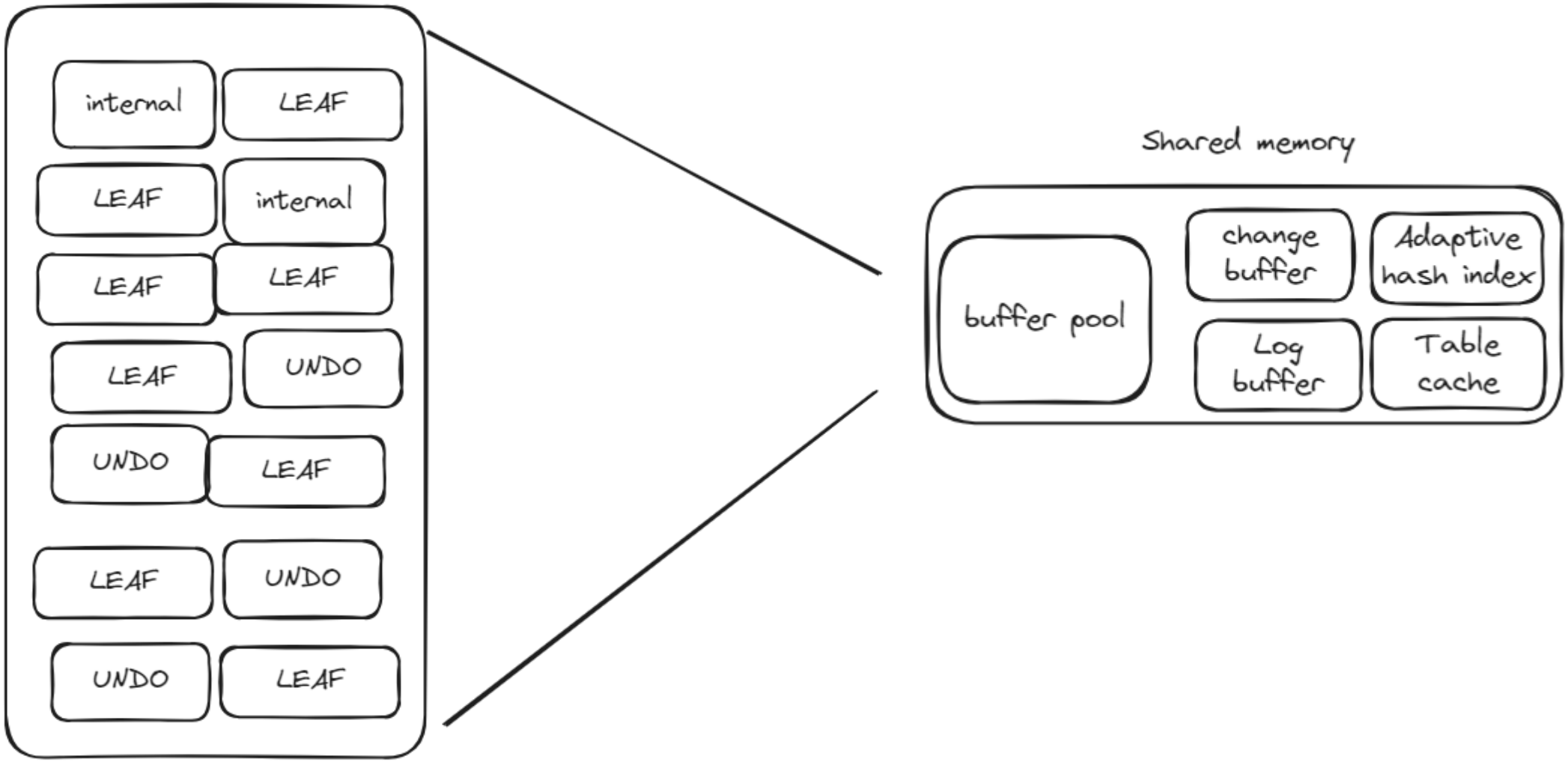
Isolation Level

1. **READ UNCOMMITTED**
2. **READ COMMITTED**
3. **REPEATABLE READ** (Default)
4. **SERIALIZABLE**

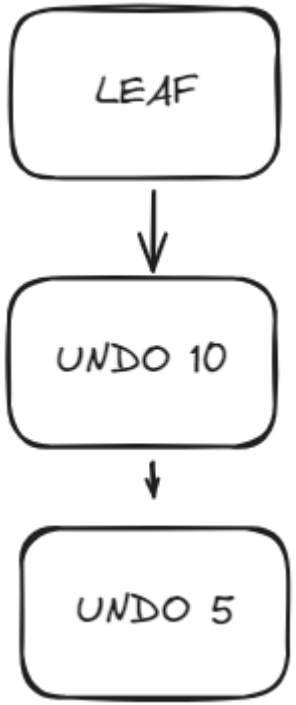
Read uncommitted



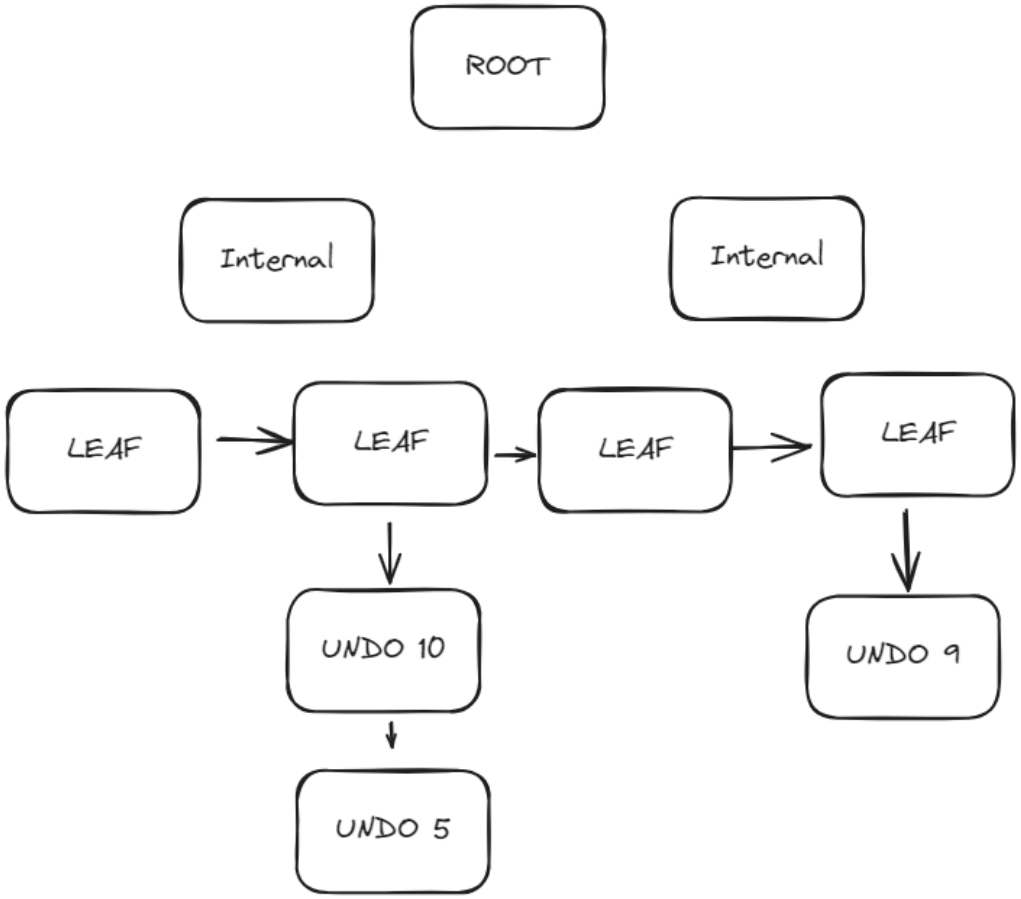
What is recorded in leaf node?



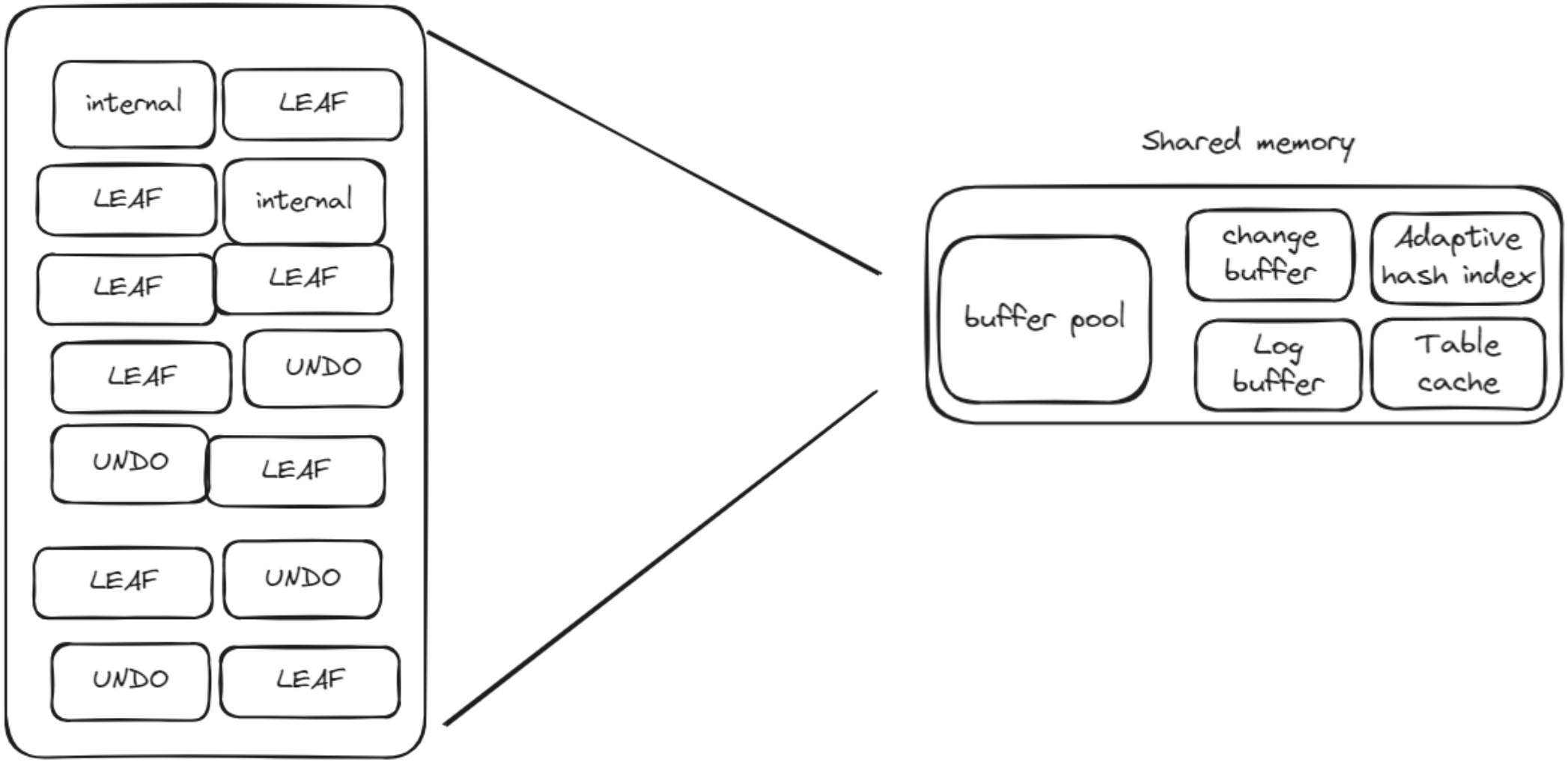
Leaf node and undo log



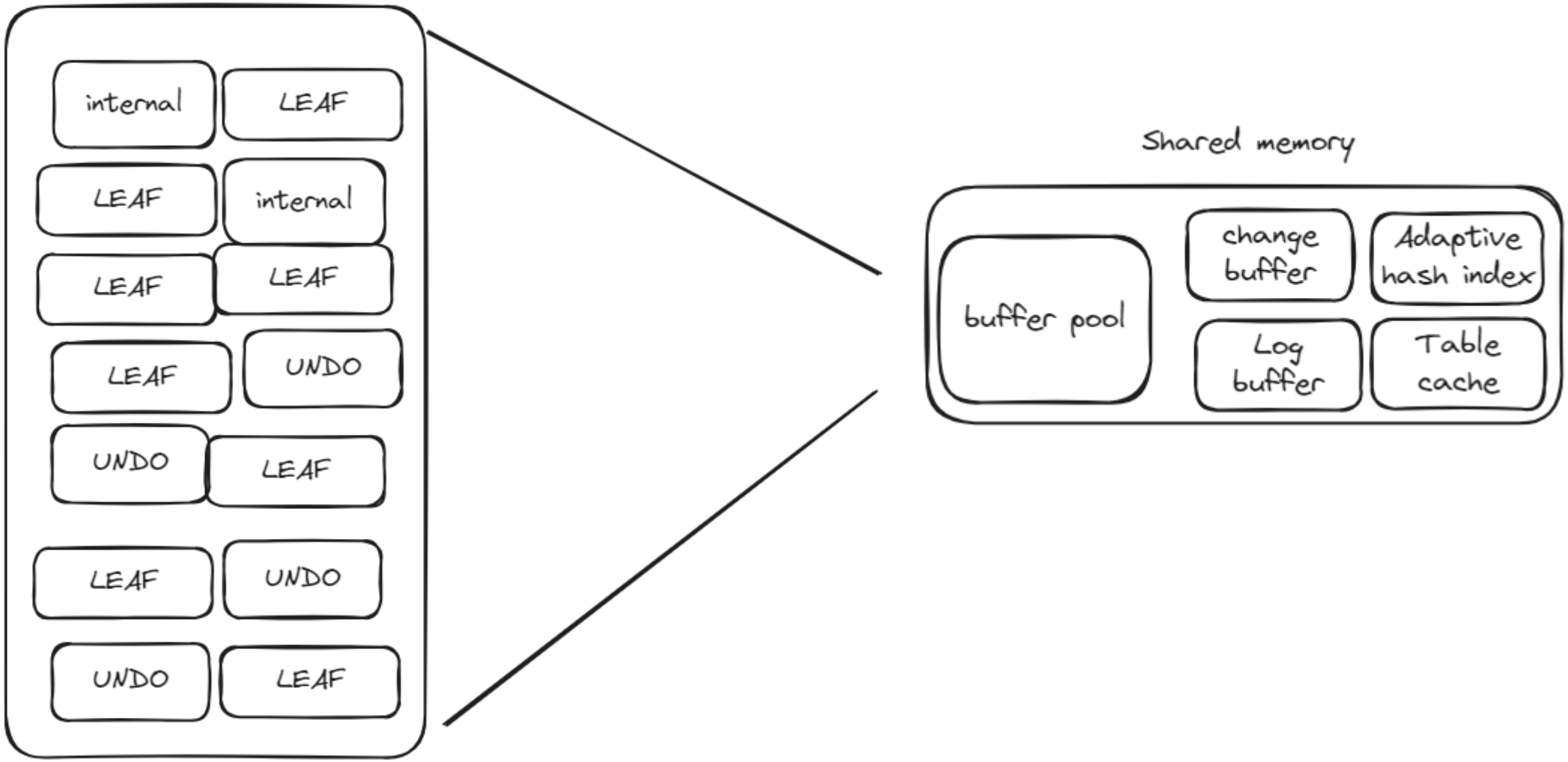
What is undo log?



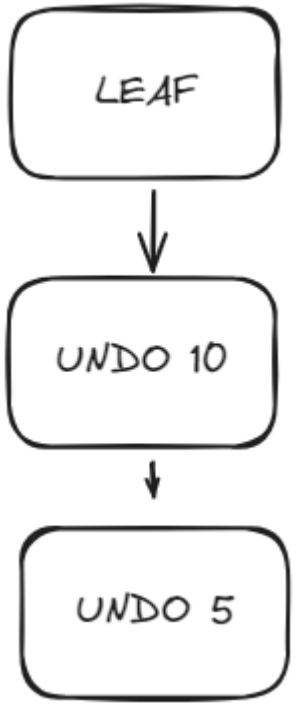
Could read uncommitted use rollback?



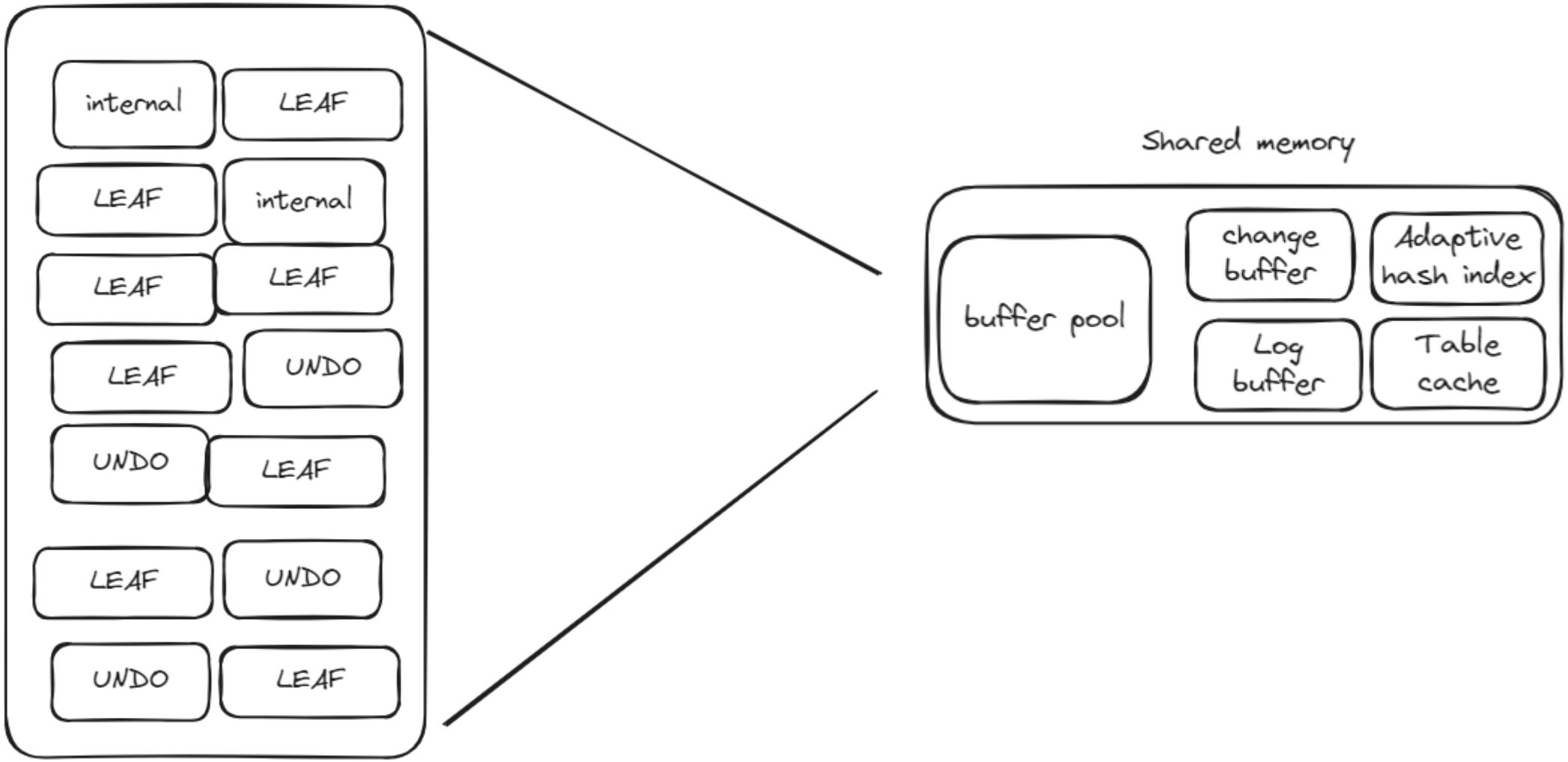
Read committed



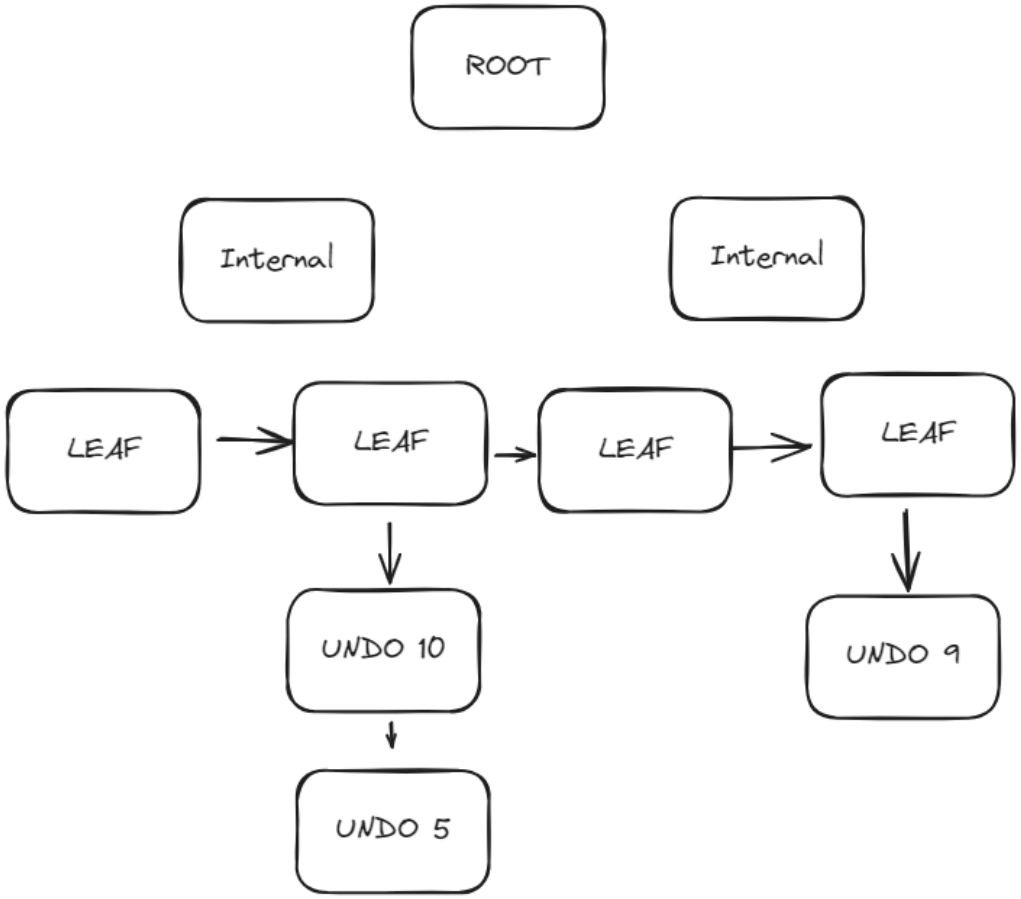
Leaf node could be uncommitted



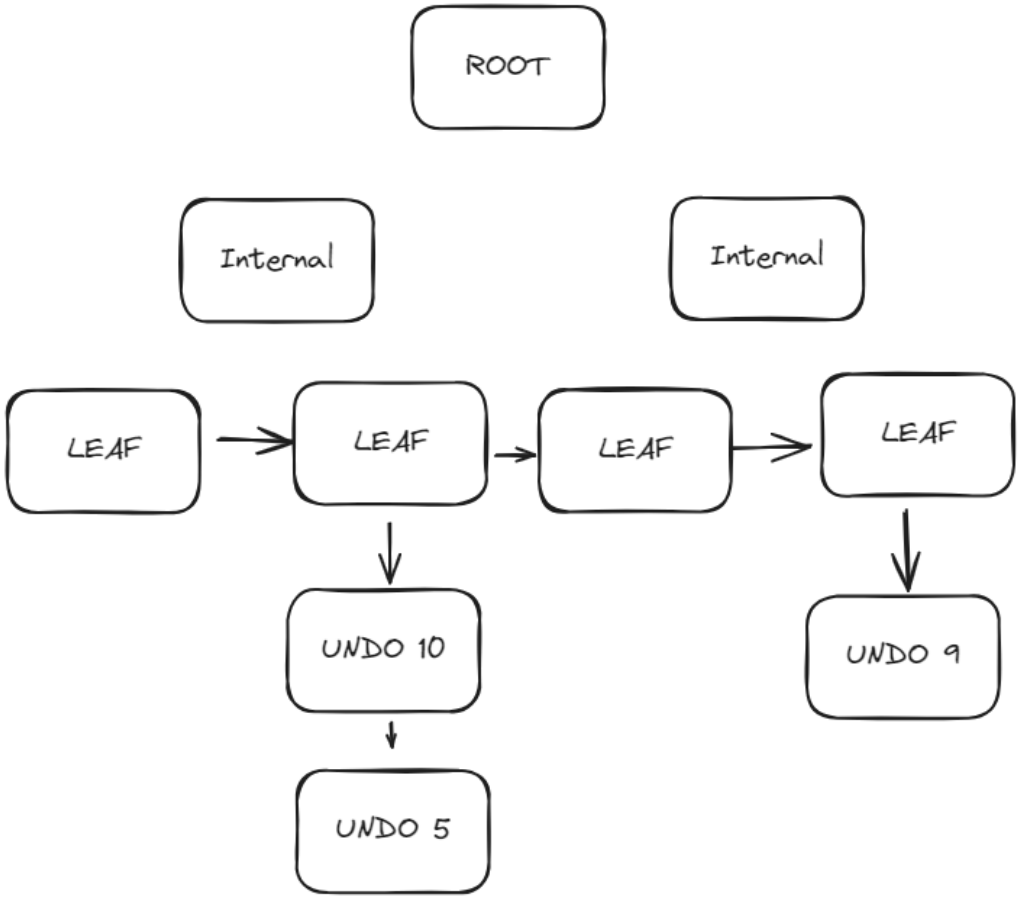
Repeatable read



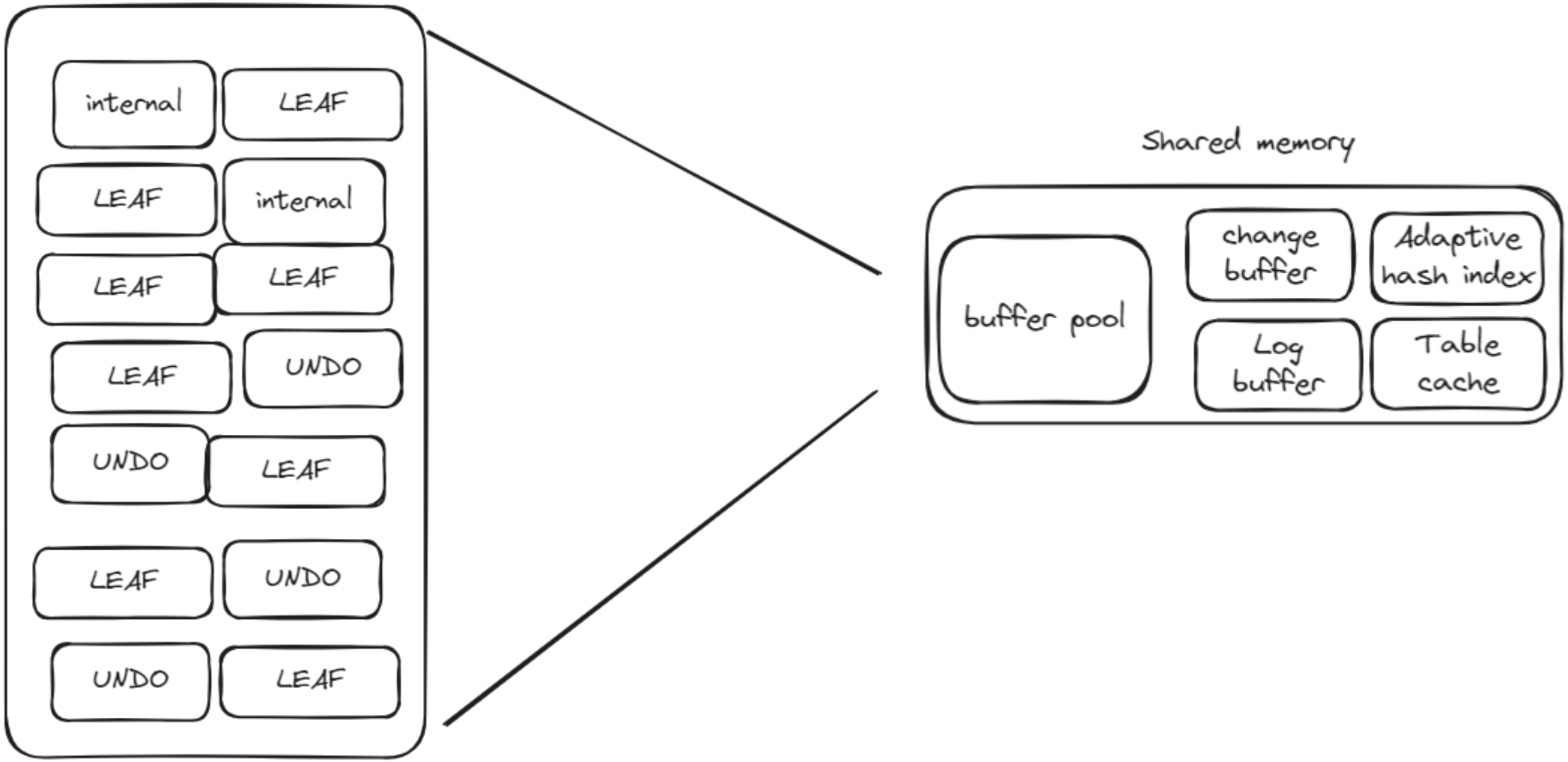
Next key lock?



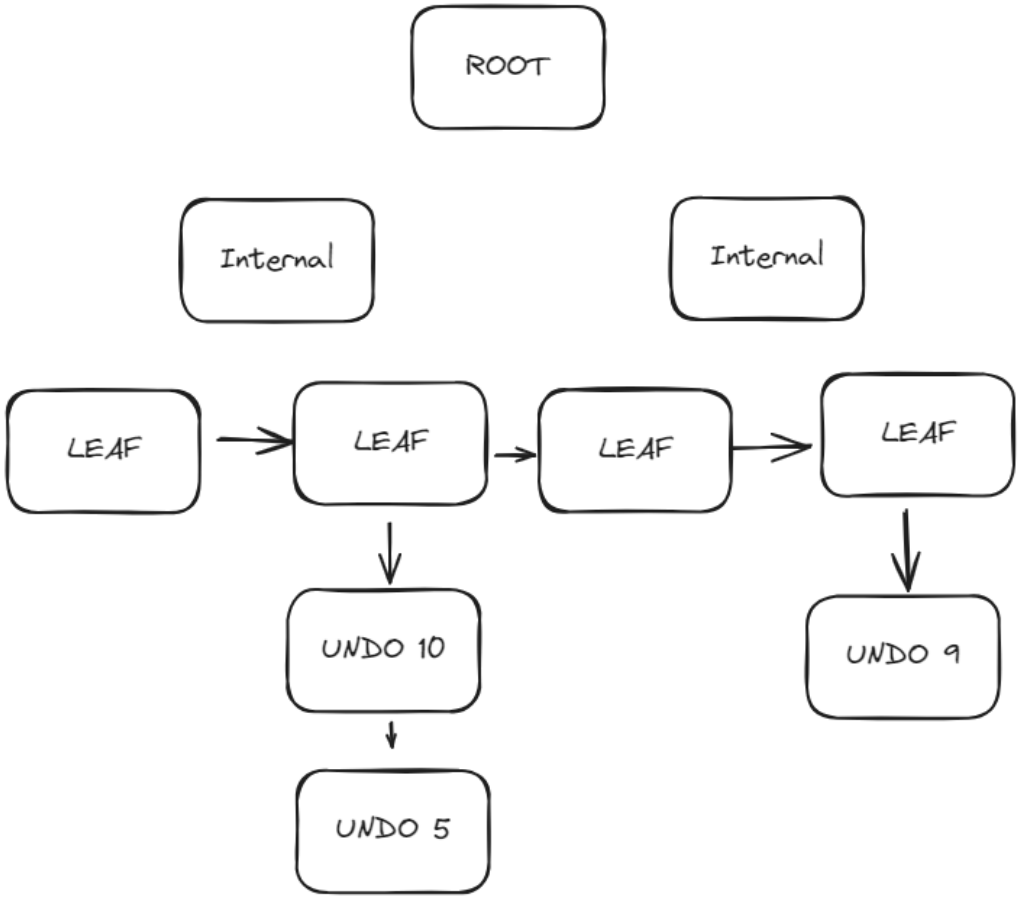
Why deadlock occurs?



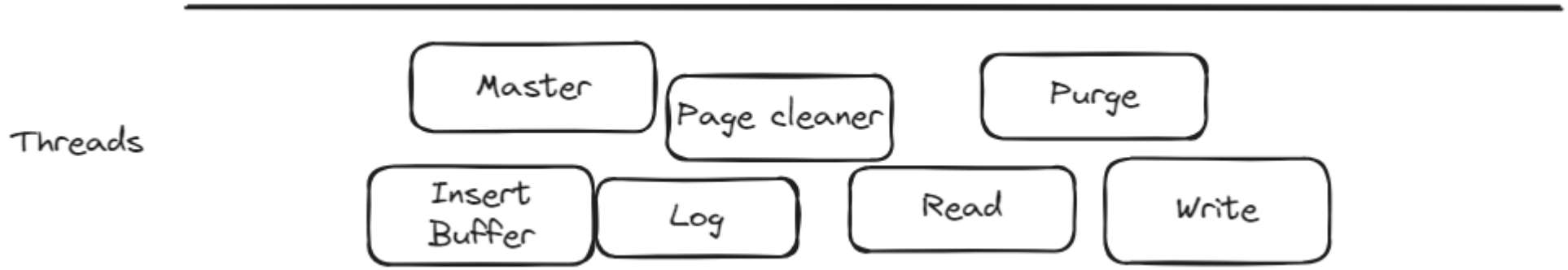
serializable



What's difference from repeatable read



Threads


Threads
Master
Page cleaner
Purge
Insert Buffer
Log
Read
Write

Storage

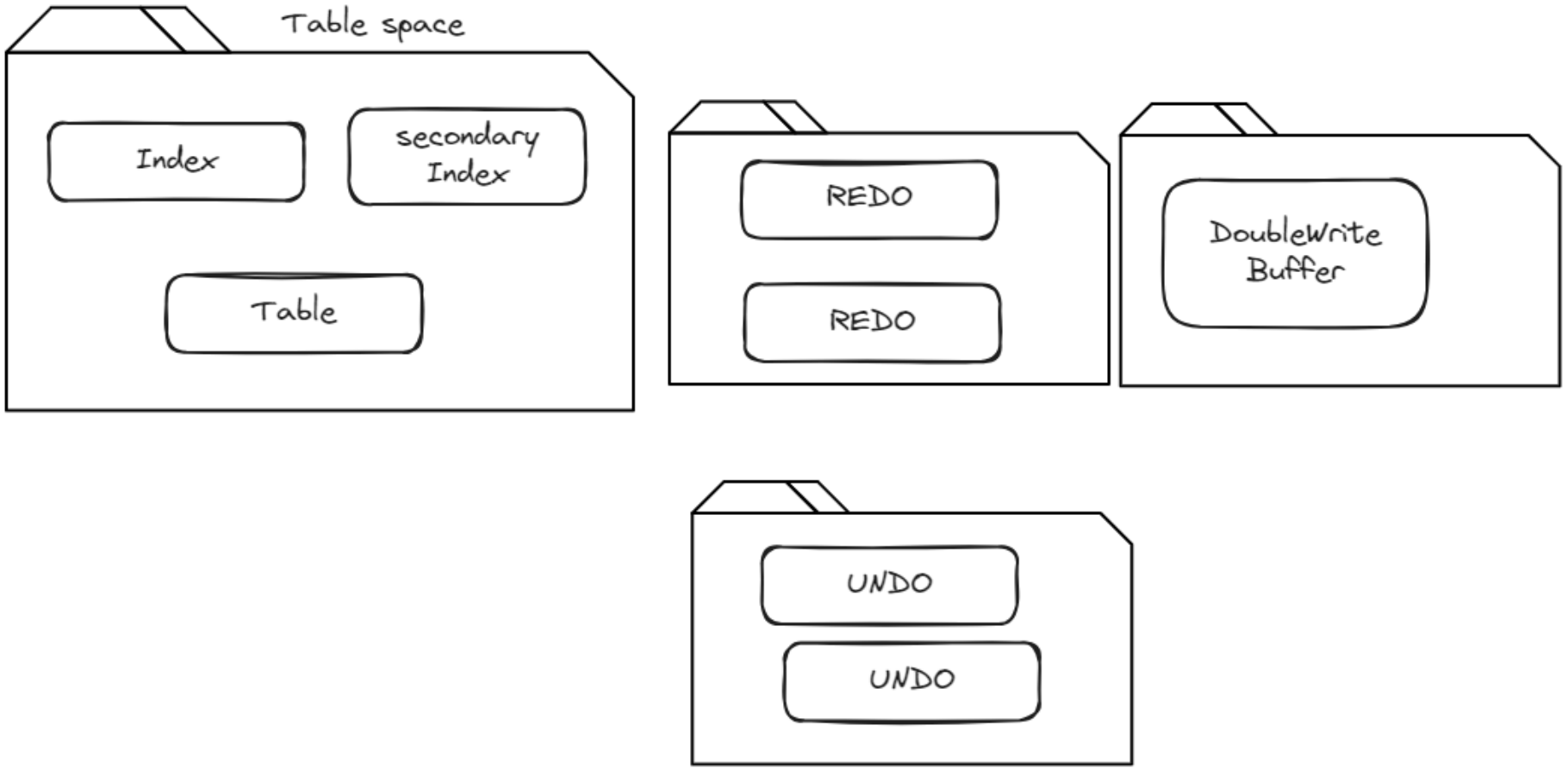
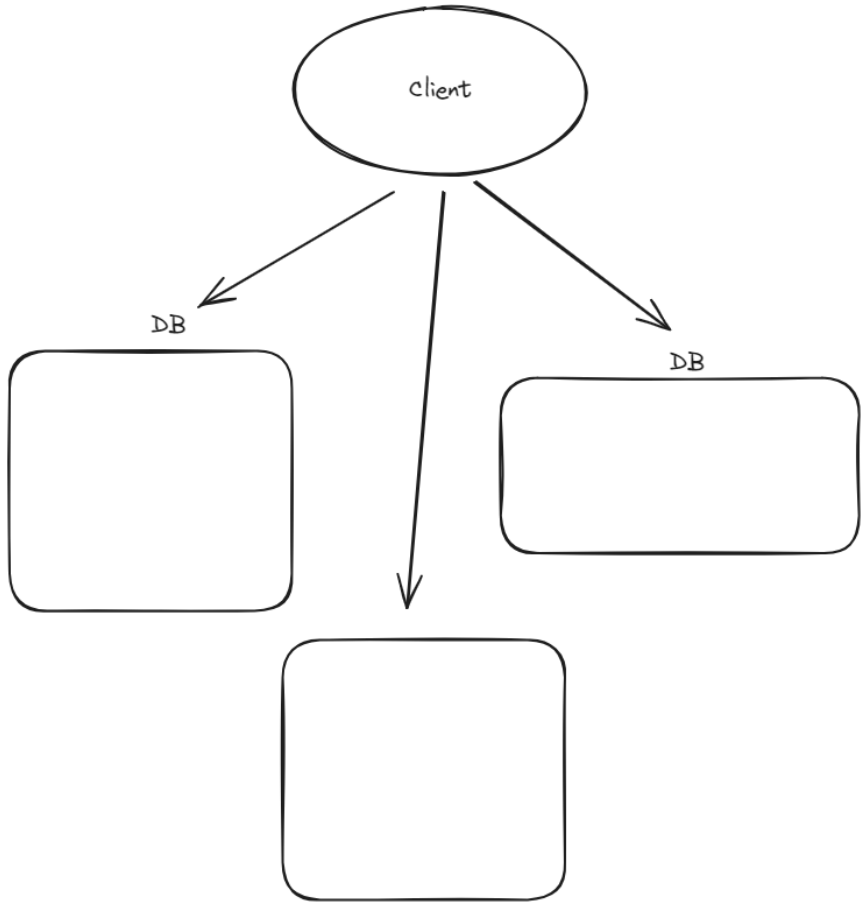
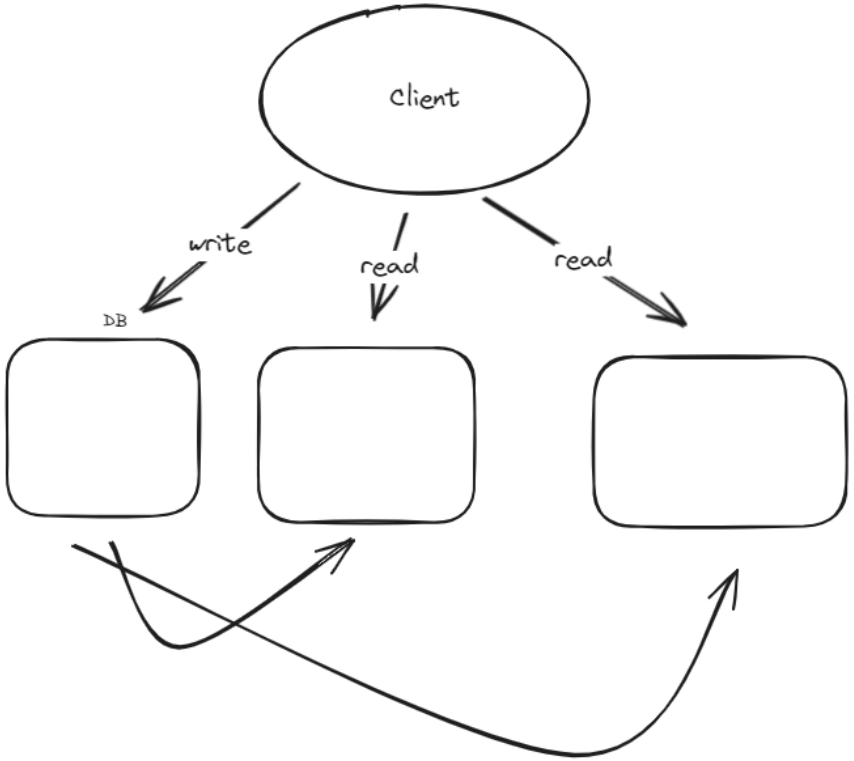


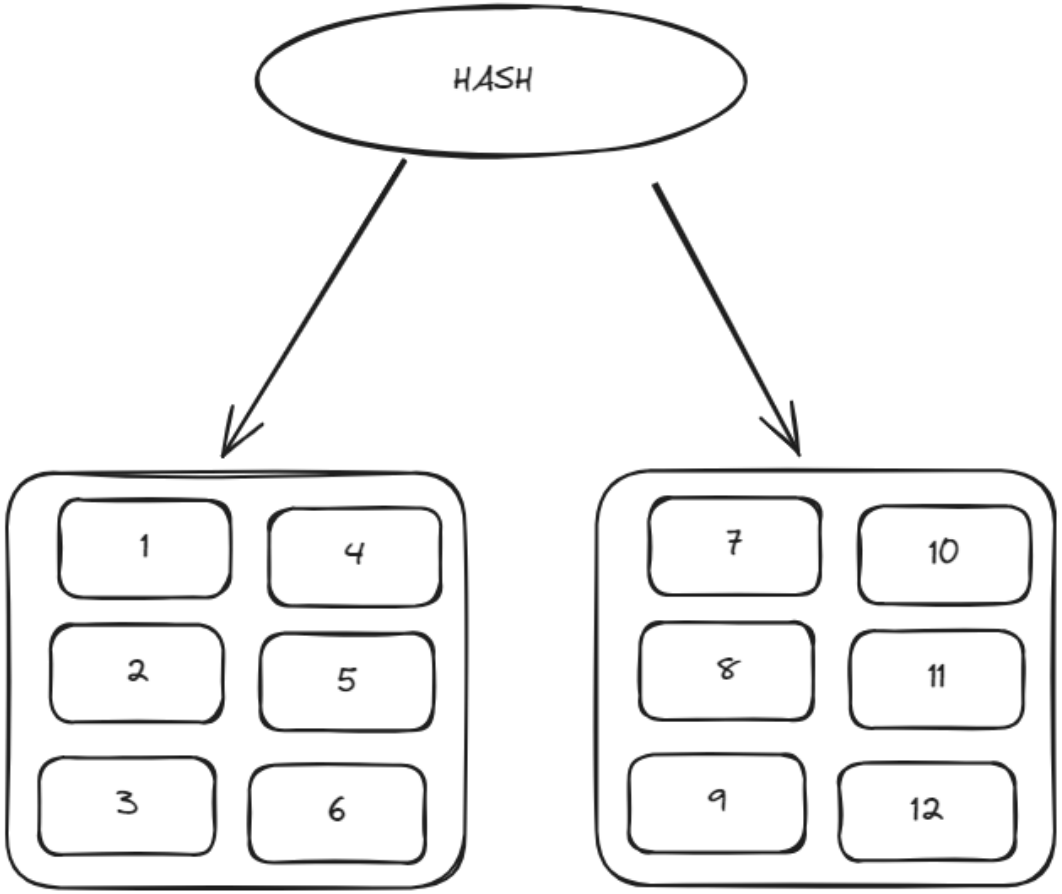
Table space
Index
secondary Index
Table
REDO
REDO
DoubleWrite Buffer
UNDO
UNDO

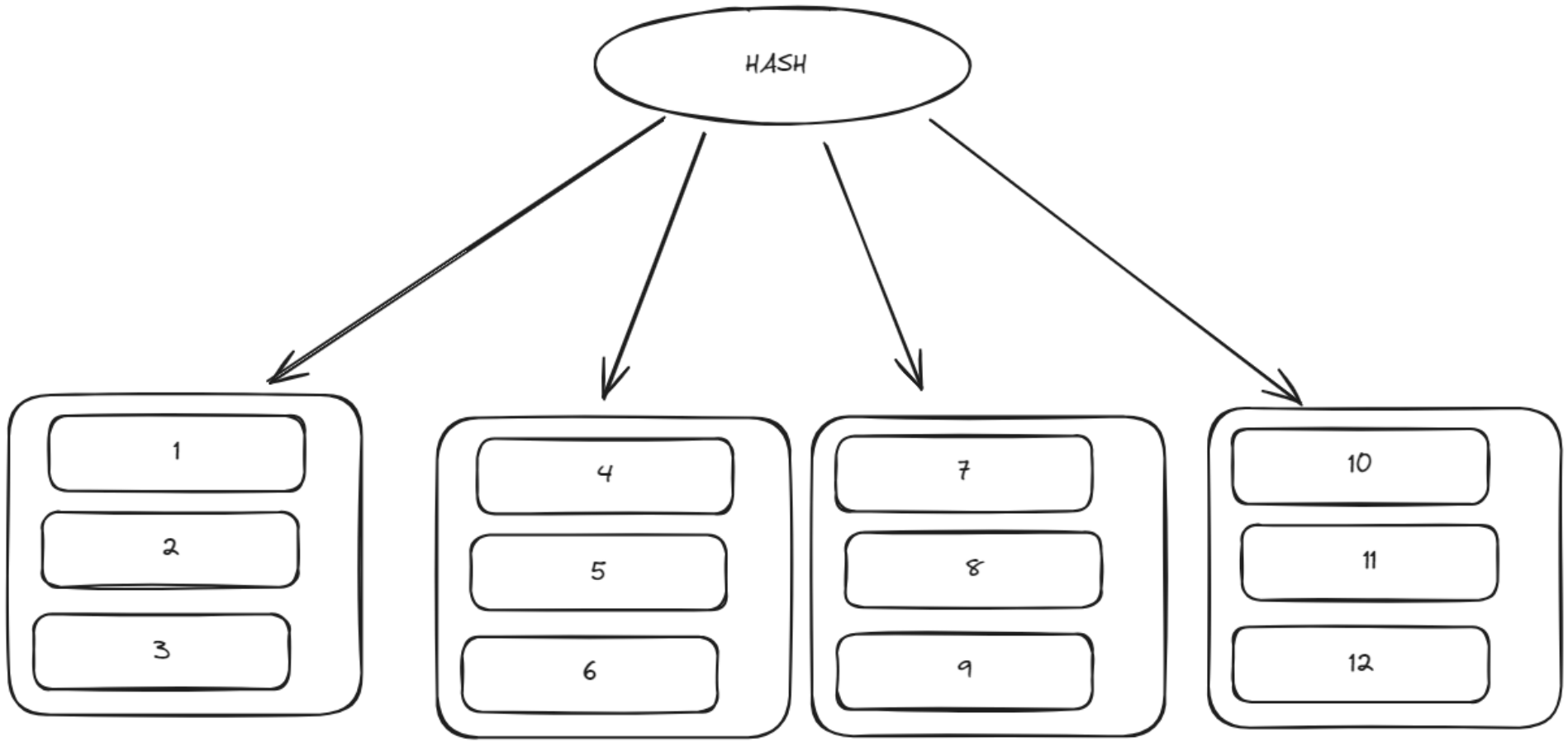
Sharding vs replication





How to make sharding scalable?




HASH
1
2
3
4
5
6
7
8
9
10
11
12