Lecture #05: Buffer Pools

https://15445.courses.cs.cmu.edu/fall2018/
Carnegie Mellon University
Prof. Andy Pavlo

1 Locks vs. Latches

Locks
- Protect the database logical contents from other transactions
- Held for transaction duration
- Need to be able to rollback changes
- Protect tuples, tables, indexes

Latches
- Protects the critical sections of the DBMS internal data structure from other threads
- Held for operation duration
- Do not need to be able to rollback changes

2 Buffer Pool

The buffer pool is an in-memory cache of pages read from disk. The DBMS always knows better so we want to manage memory and pages ourselves.

It is a region of memory organized as an array of fixed size pages. Each array entry is called a frame. When the DBMS requests a page, an exact copy is placed into one of these frames.

Metadata maintained by the buffer pool:
- Page Table: In-memory hash table that keeps track of pages that are currently in memory.
- Dirty-flag: Gets set when a thread modifies a page (will need to be written back).
- Pin Counter: Number of threads touching that page.

Optimizations:
- Multiple Buffer Pools: The DBMS can also have multiple buffer pools for different purposes. This helps reduce latch contention and improves locality.
- Pre-Fetching: The DBMS can also optimize by pre-fetching pages based on the query plan. Commonly done when accessing pages sequentially.
- Scan Sharing: Query cursors can attach to other cursors and scan pages together.

3 Replacement Policies

A replacement policy is an algorithm that the DBMS implements that makes a decision on which pages to evict from buffer pool when it needs space.

Implementation goals:
- Correctness
• Accuracy
• Speed
• Meta-data overhead

Least Recently Used (LRU)
• Maintain a timestamp of when each page was last accessed.
• DBMS picks to evict the page with the oldest timestamp.

CLOCK
Approximation of LRU without needing a separate timestamp per page.
• Each page has a reference bit
• When a page is accessed, set to 1

Organize the pages in a circular buffer with a “clock hand”
• Upon sweeping check if a page's bit is set to 1
• If yes, set to zero, if no, then evict
• Clock hand remembers position between evictions

Alternatives
Problems with LRU and Clock replacement policies:
• LRU and Clock are susceptible to sequential flooding where the buffer pool’s contents are trashed due to a sequential scan.
• It may be that the LRU page is actually important due to not tracking meta-data of how a page is used.

Better solutions:
• LRU-K: Take into account history of the last K references
• Priority hints: Allow txns to tell the buffer pool whether page is important or not
• Localization: Choose pages to evict on a per txn/query basis

4 Allocation Policies
• Global policies: How a DBMS should make decisions for all active txns.
• Local policies: Allocate frames to a specific txn without considering the behavior of concurrent txns.