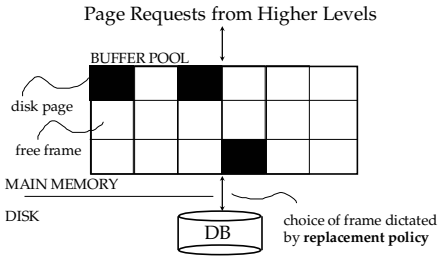
DBMS Buffer Management

Putting the data in main memory!

DBMS Buffer



- *Data must be in RAM for DBMS to operate on it!*
- *Table of <frame#, pageid> pairs is maintained.*

Requesting a Page

- If requested page is not in pool:
  - Choose a frame for *replacement*
  - If frame is dirty, write it to disk
  - Read requested page into chosen frame
- *Pin* the page and return its address.
  - *If requests can be predicted (e.g., sequential scans) pages can be pre-fetched several pages at a time!*

Managing a Page

- Requestor of page must unpin it, and indicate whether page has been modified:
  - *dirty* bit is used for this.
- Page in pool may be requested many times,
  - a *pin count* is used. A page is a candidate for replacement iff *pin count* = 0.
- CC & recovery may entail additional I/O when a frame is chosen for replacement. (*Write-Ahead Log* protocol; more later.)

Buffer Replacement Policy

- Frame is chosen for replacement by a *replacement policy:*
  - Least-recently-used (LRU), Clock, MRU etc.
- Policy can have big impact on # of I/O's; depends on the *access pattern*.
- *Sequential flooding*: Nasty situation caused by LRU + repeated sequential scans.
  - # buffer frames < # pages in file means each page request causes an I/O. MRU much better in this situation (but not in all situations, of course).

Page Replacement Algorithms

- Algorithms can be classified as
  - prefetching
  - demand
- Performance measured by

$$Miss\ rate = \frac{\#of\ \ physical\ references}{\#of\ \log ical\ references}$$

Page Replacement Algorithms (cont.)

		Age		
		No consideration	Most recent reference	First reference
references	No consideration	Random		FIFO
	Most recent reference		LRU CLOCK	
	First reference	LFU	GCLOCK	LRD

Page Replacement Algorithms (cont.)

- FIFO (First-In-First-Out)
  - replaces oldest buffer page
- LFU (Least Frequently Used)
  - replaces page with lowest reference frequency
- LRU (Least Recently Used)
  - replaces page least recently referenced / unfixed

Page Replacement Algorithms (cont.)

- CLOCK
  - simulates LRU behaviour by stepping through pages and checking use bit
- GCLOCK (Generalized CLOCK)
  - combines LFU and CLOCK
  - use bit replaced with reference counter
  - step through looking for page with 0 reference counter

Page Replacement Algorithms (cont.)

- LRD (Least Reference Density)
  - reference density relates number of references to the age of a page
  - replace page with lowest reference density

Multiple Buffer Pools

- Recent products like DB2/UDB use multiple buffer pools instead of a single buffer area
- Tablespaces assigned to specific buffer pools
- Replacements from within the one pool

DBMS vs. OS File System

OS does disk space & buffer mgmt: why not let OS manage these tasks?

- Differences in OS support: portability issues
- Some limitations, e.g., files can't span disks.
- Buffer management in DBMS requires ability to:
  - pin a page in buffer pool, force a page to disk (important for implementing CC & recovery),
  - adjust *replacement policy*, and pre-fetch pages based on access patterns in typical DB operations.