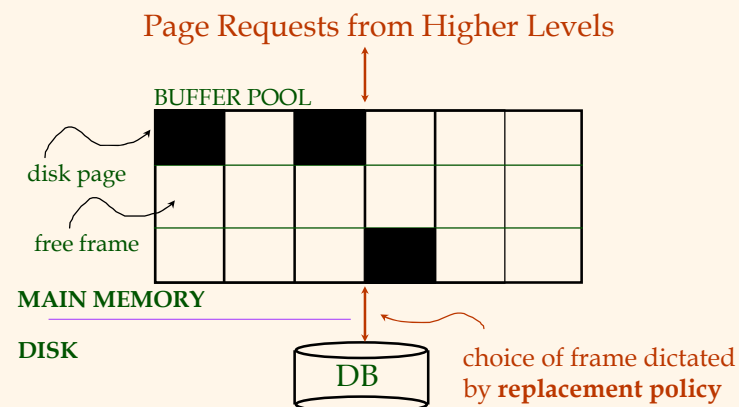


## Disk Space Management



- ❖ Lowest layer of DBMS software manages space on disk.
- ❖ Higher levels call upon this layer to:
  - allocate/de-allocate a page
  - read/write a page
- ❖ Request for a *sequence* of pages must be satisfied by allocating the pages sequentially on disk! Higher levels don't need to know how this is done, or how free space is managed.

## Buffer Management in a DBMS



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

## When a Page is Requested ...



- ❖ 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!

## More on Buffer Management



- ❖ 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).

## 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.