CONSTRUCTOR UNIVERSITY

Storing Data: Database Caching

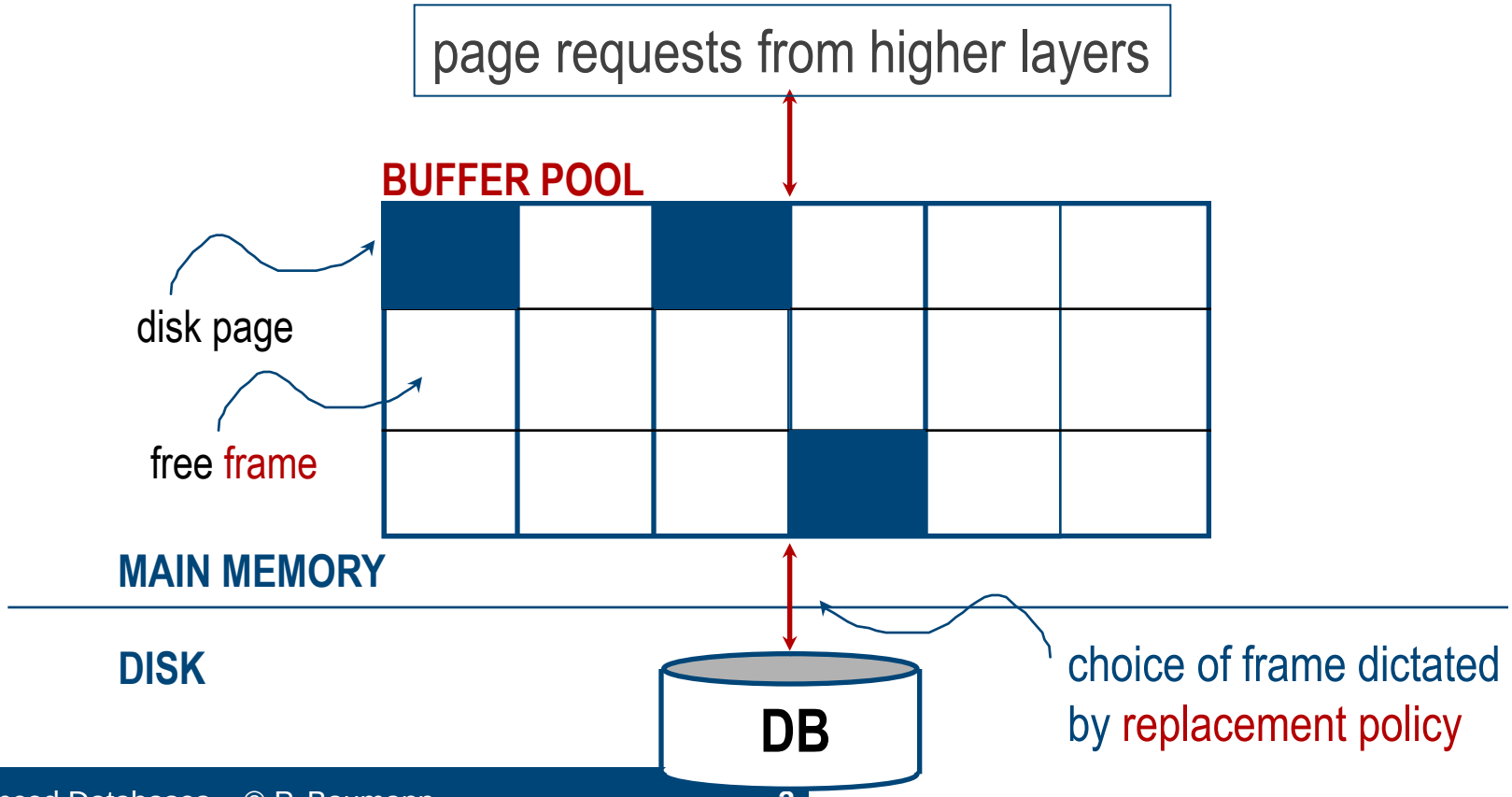
"Yea, from the table of my memory
I'll wipe away all trivial fond records."
-- Shakespeare, *Hamlet*

Disk Space Management

- Lowest layer of DBMS manages space on disk
- Higher levels call layer to:
  - allocate/de-allocate page
  - read/write page

Buffer Management in a DBMS

- Table of <frame#, pageid> pairs (plus more, see next)



When Page is Requested ...

- If page not in pool:
  - Choose frame for replacement
  - If frame dirty, write to disk
  - Read page into frame
- Pin page & return address

- *If possible, arrange blocks sequentially on disk*
  - *minimize seek and rotational delay*
- *For sequential scan (access predictable!), pre-fetching is a big win*

NB:
'page' ≈'block'

More on Buffer Management

- Page requestor must unpin
  & indicate whether page has been modified
  - *dirty* bit
- Page in pool may be requested many times
  - *pin count:* page is candidate for replacement iff *pin count* == 0
- CC & recovery: additional I/O when replacing frame
  - *Write-Ahead Log* protocol

Buffer Replacement Policy

- Frame chosen for replacement by *replacement policy*:
  - Least-recently-used (LRU), Clock, MRU etc.
- Policy can have big impact on # of I/O's; depends on *access pattern*

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 page in buffer pool, force page to disk (CC & recovery!)
  - adjust replacement policy + pre-fetch pages
    based on access patterns in typical DB operations

Summary

- Disks provide cheap, non-volatile storage
  - Random access, but cost depends on location of page on disk
  - important to arrange data sequentially to minimize seek and rotation delays
- Buffer manager brings pages into RAM
  - Page stays in RAM until released by requestor
  - Written to disk when frame chosen for replacement (which is sometime after requestor releases the page)
  - Choice of frame to replace based on replacement policy
  - Tries to pre-fetch several pages at a time