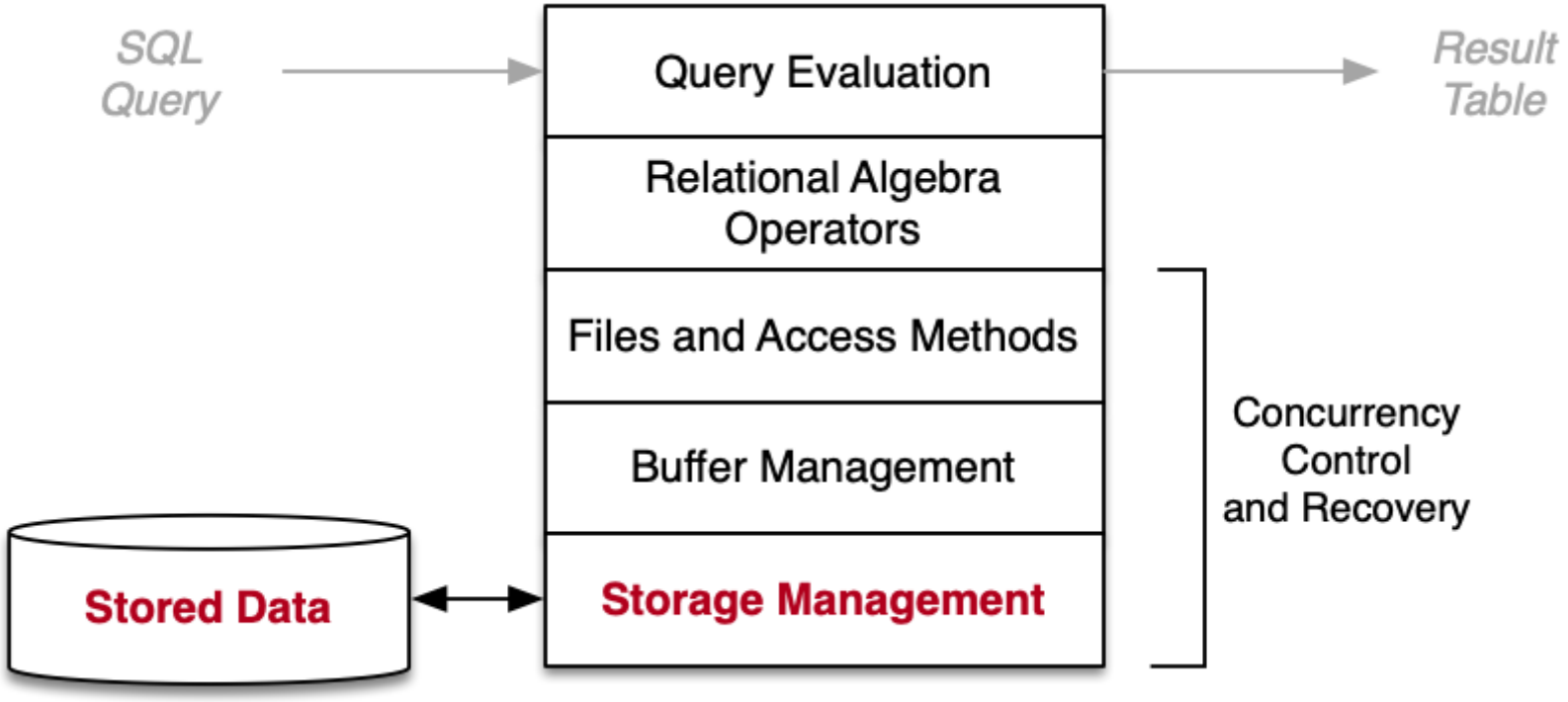
Storage Management

- Storage Management
- Storage Technology
- Views of Data in Query Evaluation
- Storage Management
- Cost Models

❖ Storage Management

Lowest levels of DBMS related to storage management:



❖ Storage Technology

Persistent storage is

- large, cheap, relatively slow, accessed in blocks
- used for long-term storage of data

Computational storage is

- small, expensive, fast, accessed by byte/word
- used for all analysis of data

Access cost HDD:RAM ≅ 100000:1, e.g.

- 10ms to read block containing two tuples
- 1μs to compare fields in two tuples

❖ Storage Technology (cont)

Hard disk drives (HDD) are well-established, cheap, high-volume, ...

- spinning magnetic medium
- access requires moving r/w head to position
- transfers blocks of data (e.g. 1KB)

Latency: move to track + spin to block = ~10ms (avg)

Volume: one HDD can store up to 20TB (typically 4TB/8TB/...)

Summary: very large, persistent, slow, block-based transfer

❖ Storage Technology (cont)

Solid state drives (SSD) are modern, high-volume devices ...

- faster than HDDs, no latency
- can read single items
- update requires block erase then write
- over time, writes "wear out" blocks
- require controllers that spread write load

Volume: one SSD can store up to 8TB (typically 1TB/2TB/...)

Summary: large, persistent, fast, (partly) block-based transfer

❖ Storage Technology (cont)

Comparison of storage device properties:

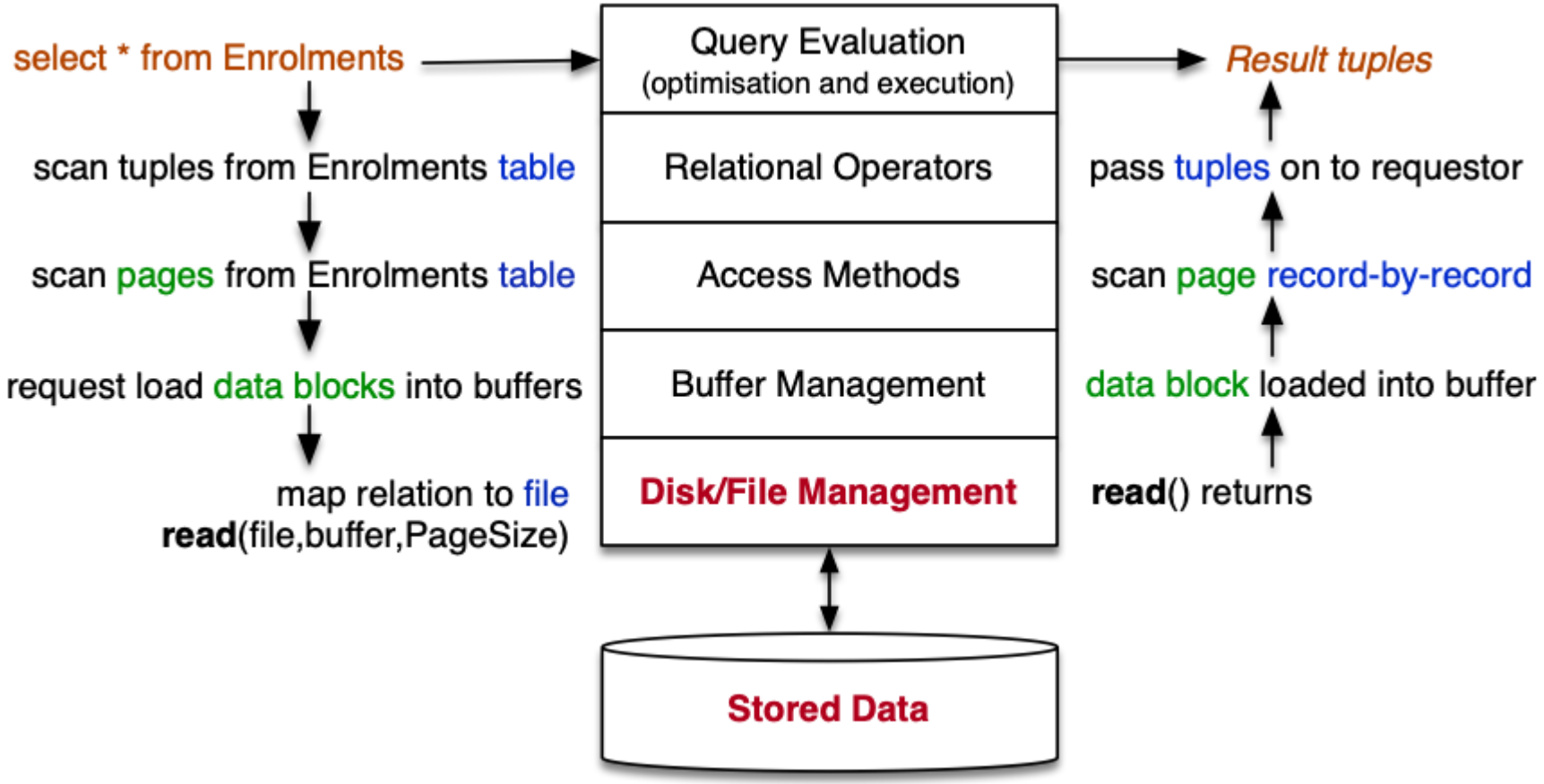
	RAM	HDD	SDD
Capacity	~ 32GB	~ 8TB	~ 2TB
Cost/byte	~ $10 / GB	~ $40 / TB	~ $200 / TB
Read latency	~ 1µs	~ 10ms	~ 50µs
Write latency	~ 1µs	~ 10ms	~ 900µs
Read unit	byte	block (e.g. 1KB)	byte
Writing	byte	write a block	write on empty block

❖ Storage Technology (cont)

Aims of storage management in DBMS:

- provide view of data as collection of pages/tuples
- map from database objects (e.g. tables) to disk files
- manage transfer of data to/from disk storage
- use buffers to minimise disk/memory transfers
- interpret loaded data as tuples/records
- basis for file structures used by access methods

❖ Views of Data in Query Evaluation



❖ Views of Data in Query Evaluation (cont)

Representing database objects during query execution:

- **`DB`** (handle on an authorised/opened database)
- **`Rel`** (handle on an opened relation)
- **`Page`** (memory buffer to hold contents of disk block)
- **`Tuple`** (memory holding data values from one tuple)

Addressing in DBMSs:

- **`PageID = FileID+Offset`** ... identifies a block of data
  - where **`Offset`** gives location of block within file
- **`TupleID = PageID+Index`** ... identifies a single tuple
  - where **`Index`** gives location of tuple within page

❖ Storage Management

Topics in storage management ...

- Disks and Files
  - performance issues and organisation of disk files
- Buffer Management
  - using caching to improve DBMS system throughput
- Tuple/Page Management
  - how tuples are represented within disk pages
- DB Object Management (Catalog)
  - how tables/views/functions/types, etc. are represented

❖ Cost Models

Throughout this course, we compare costs of DB operations

Important aspects in determining cost:

- data is always transferred to/from disk as whole blocks (pages)
- cost of manipulating tuples in memory is negligible
- overall cost determined primarily by #data-blocks read/written

Complicating factors in determining costs:

- not all page accesses require disk access (buffer pool)
- tuples typically have variable size (tuples/page ?)

More details later ...

Produced: 17 Feb 2021