

## Question 1: Relational Model ..... [305 points]

- (i) [10 points] **Anatomy of a DBMS:**  
Provide brief definitions of these components of a DBMS.
- Connection Manager
  - Query Parser
  - Query Optimizer
  - Query Executor
  - Lock Manager
  - Access Methods
  - Buffer Pool Manager
  - Log Manager
  - Disk Manager
  - Memory Manager
- (ii) [10 points] **Data Independence:**  
How does a relational DBMS support data independence?
- (iii) [10 points] **Layered Architecture:**  
List an advantage and a disadvantage of using a layered architecture.
- (iv) [10 points] **Layered Architecture:**  
List the input and output of each layer.
- Query Layer
  - Access Layer
  - Storage Layer
- (v) [10 points] **Layered Architecture:**  
Distinguish between storage and buffer management.
- (vi) [10 points] **Impact of Hardware:**  
Explain Moore's Law and its significance. Is it still valid?
- (vii) [10 points] **Storage Hierarchy:**  
List an advantage and a disadvantage of using a storage hierarchy.
- (viii) [10 points] **Locality of Reference:**  
Define: (1) spatial locality, and (2) temporal locality.
- (ix) [15 points] **Hard Disk Drive:**  
Explain the imbalance between random and sequential I/O performance in a HDD. Is this gap growing or shrinking over time?
- (x) [20 points] **Hard Disk Drive:**  
List three performance optimizations used in a DBMS to circumvent the limited random I/O throughput of HDD.

- (xi) **[10 points] Access Times:**  
List the time taken to access these elements of the storage hierarchy (round up to the closest power of 10).
- L2 Cache
  - DRAM
  - NVM
  - SSD
  - HDD
  - Network Storage
- (xii) **[10 points] Disk-Oriented DBMS:**  
Define a page table. How is it used in a DBMS?
- (xiii) **[10 points] Memory Mapping:**  
Define memory mapping. How is it used in a DBMS?
- (xiv) **[10 points] Memory Mapping:**  
What is a system call? Briefly explain the purpose of system calls.
- mmap
  - madvise
  - mlock
  - msync
- (xv) **[10 points] Memory Mapping:**  
Can the DBMS control the ordering of disk writes using only mmap?
- (xvi) **[10 points] Why not use the OS?:**  
List three benefits of using a buffer manager instead of mapping memory using the OS.
- (xvii) **[10 points] Scheduling:**  
What are the benefits of scheduling I/O operations in a storage manager?
- (xviii) **[10 points] File Layout:**  
Distinguish between the notions of a page in a DBMS, a page in an OS, and a page on disk. Why do they differ from each other?
- (xix) **[10 points] Storage Management:**  
Distinguish between these three entities: file, page, and tuple.
- (xx) **[10 points] File Layout:**  
Define these types of file organization.
- Heap file organization
  - Sequential file organization
  - Hashing file organization
  - Clustering file organization

- Log-structured file organization
- (xxi) **[10 points] File Layout:**  
Why is heap file called a "heap"?
- (xxii) **[10 points] File Layout:**  
Distinguish between these two ways to organize a heap file: linked list and page directory. List an advantage and a disadvantage of the latter approach.
- (xxiii) **[10 points] Page Layout:**  
List two benefits of using a slotted page layout.
- (xxiv) **[10 points] Log-Structured File Organization:**  
List an advantage and a disadvantage of log-structured file organization.
- (xxv) **[10 points] Log-Structured File Organization:**  
Define compaction. Why is it needed?
- (xxvi) **[10 points] Tuple Layout:**  
Do we need to store meta-data about the tuple's schema along with every tuple in a relational DBMS. Why?
- (xxvii) **[10 points] Tuple Layout:**  
Do we need to store meta-data about the tuple's schema along with every tuple in a document DBMS (non-relational). Why?
- (xxviii) **[10 points] Tuple Layout:**  
Distinguish between normalization and de-normalization. List an advantage and a disadvantage of de-normalizing tuples.
- (xxix) **[10 points] Tuple Identifier:**  
How does a disk-oriented DBMS refer to an individual tuple?