

Database Project (SWE3033) – Midterm, Fall 2022 (Instructor: Sang-Won Lee)

Student ID: _____ Name: _____ Score: _____/100

Please write the answer sheet in Korean or English

1. List at least **three** differences between OLAP and OLTP workload. [10pt]

2. During the lab experiment, we ran TPC-C benchmarks with various configurations and analyzed the DBMS performance by measuring some performance metrics. State **at least three** performance metrics and elaborate on what those metrics indicate. [10pt]

3. Which type of SQL statement only reads tuples from pages? Also, name at least one read-only TPC-C transaction. [5pt each * 2= 10pt]

4. Full table/index scans or read-ahead features can worsen the hit ratio of the buffer pool. Explain **two features** that make MySQL buffer pool scan resistant. [5pt each * 2= 10pt]

5. Assume you have to manage DB with the size of 100GB. You already have an SSD, and you need to purchase a DRAM that is 500 times faster than the SSD. The hit ratio varying the buffer size is shown in the table. What is the **ideal DRAM capacity** in terms of cost *efficiency*? Choose the DRAM capacity between 10~50GB and explain the reason why. [10pt]

DRAM SIZE	10GB	20GB	30GB	40GB	50GB
HIT RATIO	95%	97%	98%	98.5%	98.75%

6. Explain why choosing clean pages over dirty pages is better for victim selection upon page miss. (Without considering hit ratio) [10pt]

7. In PA1, we searched for the ideal LRU_scan_depth. Explain the key role of LRU_scan_depth in the page miss scenario and the page cleaner thread, respectively. [5pt * 2 = 10pt]

8. Suppose you used a hard disk (HDD) as the storage device. Then, you changed the device to a flash SSD (SSD). The performance of each device is shown in the table below. As you change the device, you notice the tendency of the victim selection method changes upon page miss (i.e., Step 1, Step 2, and Step 3). Predict the change and explain why. (Hint: The foreground user thread issues page *read* requests, while the background page cleaner thread issues page *write* requests to flush dirty pages in the LRU list tail.) [10pt]

RANDOM IOPS	READ/S (4KB)	WRITE/S (4KB)
HDD	1,400	1,300
SSD	102,000	10,000

9. Describe three main components of RocksDB and explain their characteristics. [10pt]

10. (1) How can we calculate space amplification? (2) Explain why space amplification worsens in B+tree-based storage engines (e.g., MySQL/InnoDB) compared to LSM tree-based storage engines (e.g., RocksDB). [5pt each * 2 = 10pt]