

## Database Management System (DBA) – 1 hour

### FINAL TEST

**Student ID:** \_\_\_\_\_

**Student Name:** \_\_\_\_\_

**Class:** \_\_\_\_\_

*Instruction:*

(\*) Write answer on your document and submit to submission box.

(\*) Submission's cut-off enabled: It means if you do not submit on time, box will be automatically closed without notifications.

#### **PART A. THEORY (4pts – 0.25 each)**

**1. Which one is not a characteristic of well-designed naming convention in key-value database?**

- a. Use meaningful and unambiguous naming components, such as 'cust' for customer or 'inv' for inventory.
- b. Use range-based components when you would like to retrieve ranges of values. Ranges include dates or integer counters.
- c. Use a common delimiter when appending components to make a key. The ':' is a commonly used delimiter, but any character that will not otherwise appear in the key will work.
- d. Keep keys as long as possible.

**2. Which design pattern provides some of features of relational database transactions in RDBMS**

- a. Time to live (TTL)
- b. Aggregates.
- c. Atomic aggregates.
- d. Enumerable keys

**3. What are enumerable keys**

- a. is a pattern that supports different attributes for different subtypes of an entity
- b. are keys that use counters or sequences to generate new keys

c. are sets of key-value pairs that allow for looking up keys or values by other attribute values of the same entity

**4. Which is NOT a characteristic of key-value database**

- a. Simplicity
- b. Speed
- c. Scalability
- d. Searchable value
- e. All of the above characteristics are correct

**5. Which of the following rules are NOT belongs to JSON Object in Document database**

- a. Documents consist of key-value pairs separated by commas
- b. Values can be numbers, strings, Booleans (true or false), arrays, objects, or the NULL value
- c. The values of objects are listed as key-value pairs within curly brackets, that is, { and }
- d. Keys can be any types (number, string, float, etc.)

**6. NoSQL databases is used mainly for handling large volumes of \_\_\_\_\_ data.**

- a. unstructured
- b. structured
- c. semi-structured
- d. all of the mentioned

**7. Which of the following is correct statement of Google BigTable ?**

- a. Developers have dynamic control over columns
- b. Data values are indexed by row identifier, column name, and a timestamp
- c. Reads and writes of a row are atomic
- d. Rows are maintained in a sorted order
- e. All of features above are correct
- f. None is correct.

**8. In Column-family database, Primary key in a keyspace is a \_\_\_\_\_.**

- a. Column family
- b. Row

c. Column

d. Data

**9. In Column-family database, Keyspace is similar to \_\_\_\_\_ in RDBMS**

a. Table

b. Fields

c. Primary key

d. Schema

**10. Vertex represents relationships between nodes in graph database? True | False**

**11. A bipartite graph is a graph with two distinct sets of vertices where each vertex in one set is only connected to vertices in the other set. (True | False)**

**12. Which of the following is not a Neo4J CQL command?**

a. CREATE

b. MATCH

c. DELETE

d. EXIT

**13. Joins in graph database are represented as**

a. constraint

b. Nodes

c. Relationships

d. Traversal

**14. The Elasticsearch default communication port is**

a. 9200/tcp

b. 3000/ip

c. 9200/udp

d. 3000/tcp

**15. In URI Search in ES, which of the following parameter denotes the number of hits to return in each request**

- A. from
- B. terminate\_after
- C. size**
- D. timeout

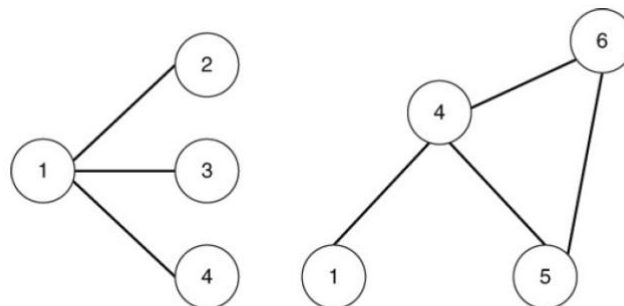
16. A \_\_\_\_\_ breakdown fields values of a document into a stream, and inverted indexes are created and updates using these values, and these stream of values are stored in the document.

- A. analyzer
- B. shard
- C. filter
- D. tokenizer**

## PART B. SHORT ANSWER (3pts – 1.5pts each)

Write short answer for these questions

1. Given the following graphs nodes



Graph A

Graph B

Figure 13.7 Two distinct graphs, A and B.

- a. Draw graph as below to show unions of 2 graphs above (0.75pts)
- b. Draw graph as below to show intersections of 2 graphs above (0.75pts)

2. List enhancements of searching using Elasticsearch over Mysql search (using LIKE pattern) (0.75pts).

When talking about advantages, you should focus on features of Elasticsearch which Mysql cannot support.

3. What is the difference between using PUT and POST in creating new document in ElasticSearch (0.75pts)

**Part C. PROBLEM SOLVING (3pts)**

1. Using Mongo, do the following tasks. (1.5pt – 0.75pts each)

Given the following model.



(\*) Note: a single book entity in the figure below can be written by many authors such as NguyenHuuCam and NguyenVanCong; therefore, using integer for author field of **Book** entity is not adequate due to incapability of using complex data structure in Visual paradigm. Note that Books entity can hold multiple authors in a single field.

(\*) Founded attribute is the year when that publisher startups.

Insert sample data to transform this model into document database using BOTH normalized and denormalized format.

2. Using MongoDB, write answer for the following queries (1.5pts)

Given the following documents in table format and collection name is employee\_info

empno	fname	lname	role	salary	departments	hiredate
1	charlie	rodgers	manager		sales, marketing	
2	sunil	chakraborty	team lead		marketing, finance	
3	sally	jones	team lead		hr, admin	
4	ben	bradley	manager		legal	
5	radha	desai	worker			
6	shruti	patel	worker			
7	mahesh	iyer	manager			

- Write query to insert document with empno = 1 and empno = 7 into mongoDB (**1 pts**)
- Using aggregation, write query to count how many of each role exists in the collection above. (**0.5pts**). Result: manager – 3, worker - 2, team lead - 2