VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

SLOT: A1 + TA1

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
FALL SEMESTER 2025-2026

Programme Name & Branch	: B.TECH. CSE
Course Code and Course Name	: BCSE406L and NoSQL Databases
Faculty Name(s)	: Dr. Joshva Devadas T, Dr. Naveenkumar J
Class Number(s)	: VL2025260101966, VL2025260101970
Date of Examination	: 17-Aug-2025
Exam Duration	: 90 minutes

Maximum Marks: 50

General instruction(s):

- Answer All Questions
- M - Max mark; CO – Course Outcome; BL – Blooms Taxonomy Level (1 – Remember, 2 – Understand, 3 – Apply, 4 – Analyse, 5 – Evaluate, 6 – Create)
- CO 1 - Define, compare and use the four types of NoSQL Databases.
  CO 2 - Explain key value databases and apply queries on those databases.
  CO 3- Explain the detailed architecture, define objects, load data, query data and performance tune Document-oriented NoSQL databases.

Q. No	Question	M	CO	BL
1.	A global logistics company is designing two new, critical, distributed services: ▪ **Package Tracking** - A system that tracks the real-time location of millions of packages. The highest priority is that customers can always access the tracking page and see the last known location of their package, even if network issues between data centers mean the location shown is a few minutes out of date. ▪ **Inventory & Customs** - A system that manages items within secure warehouses and processes customs declarations. This system must guarantee that an item cannot be in two locations at once and that a customs fee, once paid, is instantly recorded and visible across all nodes to prevent duplicate charges or release errors. Using the principles of the CAP theorem, ACID, and BASE, explain which model each of the above service aligns with and justify this alignment based on the forced trade-offs between Consistency and Availability.	10	1	3
2.	A development team is building a real-time analytics system for a popular e-commerce website. One of the requirements is to display the 10 most recently sold products across the entire site. As their first attempt using a key-value database, they design the following structure for each sales transaction: **Key:** A unique transaction ID (e.g., txn:a4e9-b1fe-c82d) **Value:** A JSON object containing product details and the time of sale (e.g., {"productId": "prod-123", "sale_time": "2023-10-26T14:30:00Z", "price": 29.99}) Is this design fundamentally flawed for efficiently meeting the "10 most recent sales" requirement? If yes explain why it is and also propose a superior key design strategy to solve this problem effectively.	10	2	4
3.	For each of the following complex system requirements, identify the single most critical design pattern for key-value databases that provides a viable solution.	10	2	3

VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

SLOT: A1 + TA1

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
FALL SEMESTER 2025-2026

	Justify why that pattern is essential to prevent system failure or extreme inefficiency. a) An API gateway must enforce a strict usage limit of 5,000 requests per user, per hour. The system must handle tens of thousands of concurrent requests and ensure that under no circumstances can a user exceed their quota, even by one request, due to race conditions. b) A large-scale cloud platform hosts data for thousands of independent projects. To comply with privacy regulations like GDPR, the platform must provide a feature to "delete all data associated with a project" immediately and completely upon request, without causing downtime or requiring a full database scan. c) A financial data service needs to store every price tick for thousands of stocks. The system must support queries that can retrieve the full price history for a single stock (e.g., AAPL) within a precise time window (e.g., from 09:30:00.000 to 09:35:00.000 on a specific date). d) An e-commerce site runs a "flash sale" where the first 100 users to click "buy" can reserve an item. The reservation is only valid for 60 seconds. If the user doesn't complete the purchase within that time, their reservation must be automatically and instantly released back into the inventory pool for others to claim. e) A photo-sharing application stores photo metadata using the photo's unique ID as the primary key. The system needs to efficiently support a query to find all photos that have been tagged with a specific keyword (e.g., "sunset")			
4.	Give the Collection: Shipments and its Documents {"shipmentId":"SHP1001","eventType":"Pickup","location":"New York, NY","timestamp":"2023-10-01T10:00:00Z","carrier":"ExpressLogi"} {"shipmentId":"SHP1002","eventType":"Pickup","location":"Los Angeles, CA","timestamp":"2023-10-01T12:30:00Z","carrier":"SpeedyShip"} {"shipmentId":"SHP1001","eventType":"In-Transit","location":"Chicago, IL","timestamp":"2023-10-02T15:00:00Z","carrier":"ExpressLogi"} {"shipmentId":"SHP1003","eventType":"Pickup","location":"New York, NY","timestamp":"2023-10-02T18:45:00Z","carrier":"ExpressLogi"} {"shipmentId":"SHP1001","eventType":"Delivered","location":"Miami, FL","timestamp":"2023-10-03T14:20:00Z","carrier":"ExpressLogi"} a) Write a query to find all events for the carrier "ExpressLogi" that are either of the type "Pickup" or occurred after 2023-10-02T12:00:00Z. b) Retrieve all events related to shipmentId "SHP1001", but structure the output to show only the eventType and location for each event, and explicitly suppress the _id field. c) Create a pipeline to determine the number of unique shipments handled by each carrier. The output should list the carrier and its count of distinct shipmentIds. d) Construct a pipeline that first counts the number of events per location, then filters the results to show only those locations with more than one event, and finally sorts the output in descending order of event count.	10	3	3

VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

REG.NO.:

SLOT: A1 + TA1

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - I
FALL SEMESTER 2025-2026

	e) Build a pipeline to find the first and most recent event timestamp for each individual shipment. The result should display the shipmentId, a firstEventTime, and a lastEventTime.			
5.	a) Illustrate with an example how the native structure of a document database fundamentally resolves the issue of Impedance mismatch.	5	3	2
	b) Compare the three primary methods of horizontal partitioning. For each, provide an ideal use case.	5		

****