Student name: Pothumulla Kankanamge Mewan Madhusha

# NoSQL Situational Quiz N1.

**NB! Grading system. Each task is graded in 0-100% percentage system, then all results are summed-up and mean value is final score for the assignment. All tasks have same weight regardless of complexity. Tasks are graded by its uniqness, insights and feasibility.**

## Task 1.

Spanish startup provides log-parser application for customers. Solution should parse terabytes of logs and send for troubleshooting only those containing error messages.

*Recommend a solution to startup given the knowledge you gained so far during the bootcamp. Think about the task from NoSQL perspective. What type of NoSQL DB would you recommend? Will it use replication or sharding? Why?*

*Fill-in the form to complete the task.*

| NoSQL DB type: | Mongo DB |
|---|---|
| Tech-stack proposed: | Language-Python, API framework – Flask API or FAST API,Log prasing-Logstash or Log4j(Need to consider also security vulnarebilities), Messege Que- AWS SQS, Frontend – typscript based front-end framework |
| Solution overview: | Setting up a log parsing application that effectively examines and troubleshoots terabytes of logs, focusing on those containing error messages, is the recommended option. MongoDB, a NoSQL database, is chosen for its scalability, flexibility, and simplicity of interaction with the technology stack of choice. |

## Task 2.

American startup is working on Dynamic Personalized Content Delivery Platform proposes log-parser application for clients. It is A platform that provides dynamic, real-time content personalization for online businesses, based on users' behaviors and preferences. By associating user identifiers (like a session ID or user ID) with personalized content or preferences, the system can provide immediate content tailoring. System should be able to assess real-time content adaptation. As soon as a user interacts with a site or app, the platform can adjust the content presented to them based on their behavior.

*Recommend a solution to startup given the knowledge you gained so far during the bootcamp. Think about the task from NoSQL perspective. What type of NoSQL DB would you recommend? Will it use replication or sharding? Why?*

*Fill-in the form to complete the task.*

| NoSQL DB type: | Mongo DB |
|---|---|
| Tech-stack proposed: | Backend - Node.js,Frontend - React.js,Log4j ,Real-time updates - WebSocket for real-time communication,Hosting - AWS (Amazon Web Services) |

| Solution overview: | MongoDB is chosen for storing user data and preferences due to its flexibility and ability to handle dynamic content and Node.js is used for server-side operations, ensuring fast real-time updates and communication with the database. In finally we can use React.js is employed for building a user-friendly interface that can seamlessly adapt to changing content based on user behavior and the reactive functions also help to Realtime changes in frontend |
| --- | --- |

## Task 3.

British IT company focuses on Comprehensive Digital Health Passport development. It is a user-centric platform that aggregates, manages, and presents a person's complete health journey. Customers will be able to track everything from doctor visits, medical history, and vaccination records, to wellness activities, exercise, and dietary habits. Health data is diverse and can come in various formats, from structured data like lab results to unstructured data like doctor's notes. Users can maintain detailed records, including family medical history, past surgeries, allergies, medications, and more. Users can share specific parts of their health passport with medical professionals, trainers, or family members. System should be able to connect with existing Electronic Health Records systems, fitness trackers, and dietary apps to pull relevant data.

*Recommend a solution to startup given the knowledge you gained so far during the bootcamp. Think about the task from NoSQL perspective. What type of NoSQL DB would you recommend? Will it use replication or sharding? Why?*

*Fill-in the form to complete the task.*

| NoSQL DB type: | Apache Cassandra |
| --- | --- |
| Tech-stack proposed: | Backend: Java (Spring Boot),Frontend: React.js,Data Integration: Apache Kafka, Hosting: AWS |
| Solution overview: | Apache Cassandra is a suitable choice due to its ability to handle large volumes of diverse and unstructured data. It provides scalability and high availability, essential for managing the comprehensive and varied health data. Java with Spring Boot is a robust backend solution, providing a stable and scalable environment for handling data processing, integration with external systems, and managing user interactions, also this kind of system needs shading and Shading is implemented in Apache Cassandra to handle the scalability requirements of the platform. As the volume of health data grows, shading distributes the data across multiple nodes, ensuring optimal performance. |

## Task 4.

ProSCoMap is Professional Skill and Connection Mapper tool, that is designed to be used by the European Commission. It is an advanced networking platform tailored for professionals and industry leaders that goes beyond the linear connection model. It visualizes complex relationships between skills, experiences, endorsements, and people to optimize career advancement, project team formation, and mentorship opportunities. Tool will help to understand how your skills, experiences, and connections intersect and relate in a vast web of professionals. System must be able to identify individuals who have complementary or desired skills you're looking to learn or collaborate with, find potential mentors or

mentees that align with career goals and skill paths. For managers or project leaders, user can input a desired skill set and find the optimal team composition within your network or organization. Finally, based on industry trends and desired career trajectory, user must be able to get recommendations on which skills to acquire next and which connections can help.

*Recommend a solution to startup given the knowledge you gained so far during the bootcamp. Think about the task from NoSQL perspective. What type of NoSQL DB would you recommend? Will it use replication or sharding? Why?*

*Fill-in the form to complete the task.*

| NoSQL DB type: | Neo 4j |
|---|---|
| Tech-stack proposed: | Backend: Python (Django or flask),Frontend: React.js,Recommendation Engine: Apache Mahout,Hosting: AWS or GCP |
| Solution overview: | Neo4j is recommended as it excels in managing and traversing complex relationships, making it ideal for a tool like ProSCoMap, which focuses on visualizing intricate connections between skills, experiences, and professionals. Instead of shading we can use replication for this system so the Neo4j provides high availability and fault tolerance through features like clustering and replication. Replication ensures that the graph data remains accessible and resilient in case of hardware failures or network issues. |

## Task 5.

Japanese startup plans to build "BankChain" application that is Global Personalized Banking Experience System. Startup is providing customers a unified banking experience irrespective of their location. offers real-time access to accounts, personalized financial products tailored to global lifestyles, and a powerful fraud detection system that spans across continents. It's banking, reimagined for the global citizen. The dynamism of global transactions necessitates a system that can handle vast amounts of diverse data at lightning speed. Application must ensure high-speed localized access to data while maintaining a unified global view. For a user, this means instant access to their account whether they're transacting in Paris, New York, or Tokyo. Regardless of where deposits or transactions are made, users have a single view of their account, updated in real-time. By analyzing transaction patterns across different regions, the system must be able predict and identify potential fraudulent activities on a global scale. Lastly, it is planned to provide customers with insights on tax implications, investment opportunities, and savings based on a multi-region perspective.

*Recommend a solution to startup given the knowledge you gained so far during the bootcamp. Think about the task from NoSQL perspective. What type of NoSQL DB would you recommend? Will it use replication or sharding? Why?*

*Fill-in the form to complete the task.*

| NoSQL DB type: | Apache Cassandra, Redis |
|---|---|
| Tech-stack proposed: | Backend: Java (Spring Boot),Frontend: React.js,Distributed Cache: Redis,NoSQL Database: Apache Cassandra,Fraud Detection: Apache Flink,Hosting: AWS or Azure |

| | |
|---|---|
| Solution overview: | Apache Cassandra is chosen for its ability to handle massive amounts of data and provide high availability, Java with Spring Boot is a reliable choice for the backend, providing a robust framework for handling global transactions, ensuring data consistency, and integrating with Apache Cassandra, also we can use Redis as a distributed cache to enhance data retrieval speed and reduce latency. Apart from that when thinking security side of the application itself we can use Apache flink for fraud detection scenarios. Finally, Apache Cassandra's replication ensures data availability, while the distributed cache and stream processing components contribute to the high-speed and real-time nature of the BankChain application |