Iteration 1: The Foundation

Scenario: The organization is an international news agency that gathers and analyzes vast amounts of multilingual text data from various sources. They have a globally distributed team of journalists and analysts who need access to the database. Their current database setup, a traditional relational database, is not scaling well with their growing data volume and increasing access demands from different geographic locations. They need a database solution that can provide high performance, high availability, and durability, along with advanced text search capabilities for their Java-based application.

Challenges: (Total marks: 10)
- For an international news agency dealing with high volumes of multilingual text data and requiring advanced text search capabilities for their Java-based application, which Amazon RDS database engine would you choose and why? Consider the need for high availability, scalability, and distributed access in your answer. (3 marks)
- Given that the initial user base is around 50,000 with a data inflow of approximately 50GB/day, and you anticipate the user base will grow to approximately 200,000 within a year, with an estimated data inflow growth rate of 10% per quarter. Additionally, the application is read-intensive with frequent complex queries. What instance type and storage capacity would you select for the above selected RDS database at launch, and how would you plan for the projected growth? (3 marks)
- How would you devise a backup and recovery strategy capable of restoring at least a week's worth of data? (2 marks)
- The organization requires a secure architecture that provides data encryption and access control. What would be your strategy? (2 marks)

Iteration 2: Network Fortress

Scenario: With the basic database infrastructure in place, the news agency now wants to ensure the robustness and security of its network. The agency seeks a network configuration that optimally balances accessibility and security. For their reporting work, they've started running certain tasks on EC2 instances, which require secure internet access.

Challenges: (Total marks: 8)
- Design a Virtual Private Cloud (VPC) for the agency's AWS infrastructure. Specifically, how many subnets would you create, in which Availability Zones (AZs), and why? (2 marks)
- How would you configure Security Groups for your RDS instances and other EC2 instances that need access to the internet, considering both inbound and outbound rules? (2 marks)

- Discuss the rules you would set for Network Access Control Lists (NACLs) at the subnet level to provide an additional layer of security. How would these rules differ for public and private subnets? (2 marks)
- As a proactive measure, the news agency wants to minimize the impact of a potential Distributed Denial of Service (DDoS) attack. Suggest a mitigation strategy using AWS services. How would this protect your VPC and the resources within it? (2 marks)

Iteration 3: The Battle Against Complexity

Scenario: A major world event leads to a sudden surge in user demand, with the number of users expected to triple, and a ten-fold increase in data inflow rate.

Challenges: (Total marks: 8)
- In light of a major world event, user demand and data inflow are projected to increase significantly (triple the user base and a ten-fold increase in data inflow). How would you modify the architecture to accommodate this surge? (2 marks)
- How would you utilize RDS Performance Insights to monitor your database performance and identify bottlenecks? (2 marks)
- When you design your RDS database architecture, how would you ensure data integrity and consistency? (4 marks)

Iteration 4: The Final Conquest

Scenario: A hurricane threatens the location of one of your major data centers. Your architecture must withstand the catastrophe and provide reliable disaster recovery capabilities.

Challenges: (Total marks: 4)
- A hurricane threatens the location of one of your major data centers. Create a disaster recovery plan that includes replication of the RDS database across regions. How would you incorporate your plan on this? (2 marks)
- There is a need to develop a failover mechanism using AWS RDS Multi-AZ deployments to ensure minimal downtime during disaster recovery scenarios. How does your mechanism guarantee minimal service disruption? (2 marks)

How to do this assignment?
To complete this assignment, start by thoroughly reading and understanding the scenarios and challenges presented in each iteration. Write a detailed response for each challenge, explaining how your proposed architectural decisions cater to the challenge's needs and your rationale behind those decisions. Although you don't have to be exact with calculations and numbers, you should demonstrate an understanding of how to approach them and be on the right path towards the right calculations and numbers. If necessary, incorporate diagrams to visualize your architecture and data flow (Not mandatory). Once you have completed all iterations, compile your responses and diagrams into a single PDF file. Submit this PDF on Brightspace by the due date.

Academic Standards:
The assignment must adhere to academic standards, including proper citation and referencing of sources, adherence to academic writing standards, and originality of work. Plagiarism or any form of academic misconduct is not tolerated and may result in an academic integrity offense. The report must demonstrate a thorough understanding of the concepts and storage services used and must be original work. It should be well-organized, clear, concise, and free of errors in grammar and spelling.

FAQ:
Q: How many pages should the report be?
A: There is no specific page requirement, but the report should be long enough to address all the assignment questions thoroughly and clearly.

Q: Can I use external sources to support my report?
A: Yes, you may use external sources such as research papers, articles, or official documentation from cloud providers to support your report. However, you should ensure that all sources are properly cited and referenced in the report to avoid plagiarism.

Q: What citation and reference format should I use for the assignment?
A: You should use either the APA or IEEE citation and reference format.

Q: Is there any specific format for the assignment report?
A: No, there is no specific format, but make sure that your report is well-organized, properly formatted, and easy to read and understand, as this will help the marker to evaluate your work more effectively.

For any other questions or concerns, reach out to the course TA.

Good luck :)