Cloud/Data/Solution Architect - Assignment

1. Current system description:

We, a wholesaler, who have many subsidiaries in US, EU and Asia. We have concrete, distributed systems on premise in each subsidiary. Some small subsidiaries use the same systems. Current data is stored in Postgres, Oracle, Sybase servers (about 30 server instances, almost our server versions are older than 11.xx) with billion data rows and Petabyte files (CSV, MS Excel…) including data and reports. Our data is not well managed, it keeps growing to huge volume as lots of obsolete data has not been removed. Sybase servers have not got patched for a long time as support contracts had been expired and we do not have any plan to renew them. In terms of reporting, some use Power BI, some use Tableau, some use MS Excel. These servers are installed and managed in data center locally by subsidiary IT teams. Each subsidiary has its own security, RBAC mechanism. Periodically, reports of each subsidiary are generated and send to Headquarter Office (hereinafter HO) via Email or SFTP. Subsidiary managers can access to managerial information of HO. IT teams are running about 100 in-house developed applications (Java-based, .NET based) and about 10 COTS applications (including Case360, S4Hana).

2. TO-BE system expectations:

We intend to migrate our IT systems to Cloud-based infrastructure, modernize our applications and build a Data Analytics Platform (hereinafter DAP), a central platform enabling its data management strategy. The new systems need to be scalable and realiable to support our growth and disaster recovery in case needed. The new data platform integrates seamlessly diverse data sources into a unified and coherent data pool. It ensures a standardized and well-governed approach to data, serving as the authoritative source for accurate and reliable information.

1. Cloud-base infrastructure: Migrate all data, applications to Cloud (AWS/Azure/GCP). Retire or replace applications if need.
2. Compliance and Risk Mitigation: Implementing the DAP with robust governance features support compliance with data protection regulations, such as PDPA. It mitigates the risk of legal issues and financial penalties associated with data mishandling. Respect GDPR, PIPL and CCPA.
3. Enhanced Data Quality: Implements robust data governance practices, ensuring high-quality data, which improves the reliability of business analytics, reporting, and strategic planning.

4. Cross-Functional Collaboration: It promotes a culture of collaboration, allowing various business units to leverage shared insights for mutual benefit.
5. Cost Savings: Eliminating redundancies in data storage, reducing expensive integrations of multiple separate solutions, and optimizing data processing workflows. Dissolve/reduce local IT teams, transfer their responsibilities to HO IT Team.
6. Faster Time-to-Insight: We expect to derive insights more quickly with streamlined data processing and integrated analytics. This agility is crucial in responding promptly to business changes and gaining a competitive edge.
7. Adaptability to Changing Business Needs: Adapt to changing business requirements and data landscapes. This adaptability ensures that we remain agile and responsive in the face of industry shifts and technological advancements.
8. Empowered Business Users: Self-service analytics capabilities provided by the DAP empower business users to access and analyze data without heavy reliance on IT. This democratization of data leads to a more agile and responsive organization.
9. Efficient Resource Utilization: Optimize resource utilization by efficiently managing and utilizing data storage, processing power, and analytical tools. This results in a more cost-effective and resource-efficient operation.
10. AI-powered system: Generative AI should be integrated into our system in order to support our self-analytic team utilizing the power of AI.

Assignments:

As the requirements for the position, you are tasked with following:
- Create a technical proposal which includes but not limitted to:
    - Requirement Analysis.
    - Key question list (if needed) in order to make the solution.

    - Approach, solution, system architecture design for:
        o Application modernization & migration.
        o Data migration & Governance.
    - Describe what you would look at during the discovery phase. What aspects of the organization (e.g., business processes, IT systems, data, customer engagement) would you assess to design a detailed solution if needed.
    - General implementation plan with team structure, timeline and your position, responsibilities in the team.

Max 20 slides for 30-min-presentation session, then 20 min for Q&A session.

Tips:

- Feel free to make assumptions to illustrate your answers to the above questions.
- The number of slides is indicative only.
- You may leverage materials publicly available online to answer questions.