BYTEWISE FELLOWSHIP

Saddar U Din Babar

Architecture for cloud based Data Management System:

Cloud Platform:

I suggest utilizing Microsoft's Azure cloud platform for this project. Azure provides a comprehensive set of services and functionalities that perfectly match the client's needs for scalability, dependability, protection, and data handling capabilities.

Comprehensive End To End Solution:

Data Storage:

Azure provides a variety of storage services suitable for diverse data types and use cases. Considering the specific requirements of the social media app, I suggest utilizing the following services:

- ***Azure Blob Storage:*** This service is ideal for storing unstructured data like images, videos, and user-generated content.
- ***Azure Cosmos DB:*** It is recommended for storing structured data, including user profiles and social media interactions. Cosmos DB offers global distribution, automatic scaling, and low latency, ensuring excellent availability and performance.
- ***Azure SQL Database:*** For relational data storage and efficient querying, Azure SQL Database is the recommended choice. It provides data security, automatic backups, and high availability features.

ETL Process:

When it comes to ETL processes, Azure offers multiple options for implementation. For this particular project, I recommend utilizing the following services:

Azure Data Factory: It is a comprehensive data integration service that allows for the orchestration and automation of data pipelines. This fully managed service supports data

extraction from diverse sources, transformation, and loading into target storage or analytical systems.

Azure Databricks: This collaborative analytics service is based on Apache Spark and is designed for efficient processing of big data. It enables scalable and streamlined ETL workflows.

Data Processes and Analysis:

Azure provides a range of services for data processing, analysis, and machine learning, including:

Azure Synapse Analytics: This unified analytics platform combines data warehousing, big data processing, and machine learning capabilities into a single solution.

Azure HD Insight: It is a managed service that supports Apache Hadoop and Spark, allowing for large-scale data processing, exploration, and analytics.

Azure Machine Learning: This cloud-based service enables users to build, deploy, and manage machine learning models efficiently.

Security Measures:

To guarantee data security and compliance, you can utilize the following Azure services:

Azure Active Directory (Azure AD): This solution offers a complete identity and access management system for managing user identities, authentication, and authorization.

Azure Key Vault: It is a secure service for managing encryption keys, sensitive data, and credentials, ensuring their protection.

Azure Security Center: This service provides advanced threat detection and prevention capabilities, acting as a unified security management and threat protection solution.

BYTEWISE FELLOWSHIP

Saddar U Din Babar

Visualization Tools:

Azure provides a variety of choices for data visualization and reporting:

Power BI: This platform is designed for business intelligence and enables the creation of interactive dashboards, reports, and visualizations.

Azure Synapse Studio: It serves as an integrated development environment (IDE) with features for data exploration, visualization, and collaboration.

BYTEWISE FELLOWSHIP

Saddar U Din Babar

Reasons Behind Selecting The Recommended Services:

The selected Azure services possess specific attributes and advantages that align perfectly with the client's requirements:

Azure Blob Storage: It provides scalable and cost-effective storage for unstructured data, ensuring high availability and durability.

Azure Cosmos DB: With global distribution, low latency, automatic scaling, and multi-model capabilities, it allows seamless storage and retrieval of social media interactions.

Azure SQL Database: This fully managed relational database service offers built-in intelligence, security, and high availability.

Azure Data Factory: It simplifies the orchestration and automation of data pipelines, supporting various data sources and destinations.

Azure Databricks: It enables efficient and scalable big data processing, including data transformation and exploration.

Azure Synapse Analytics: This service integrates data warehousing, big data processing, and machine learning capabilities, enabling end-to-end analytics.

Azure HD Insight: As a managed Apache Hadoop and Spark service, it facilitates large-scale data processing and analytics.

Azure Machine Learning: It streamlines the development, deployment, and management of machine learning models.

Azure Active Directory: It ensures secure identity management, single sign-on, and access control for the application.

Azure Key Vault: It safeguards sensitive data, encryption keys, and credentials, providing secure access and management.

Azure Security Center: This service offers advanced threat detection, monitoring, and compliance management.

Power BI: It enables interactive data visualization and reporting, empowering users to derive insights from the application's data.

Azure Synapse Studio: It provides a unified development environment for data exploration, visualization, and collaboration.

Cost Analysis and Project Duration:

The cost of the project will vary depending on factors such as data volume, usage patterns, and the selected Azure services. To accurately estimate the costs, a detailed analysis considering the following aspects should be conducted:

Storage costs: The pricing will depend on the specific storage services utilized (e.g., Blob Storage, Cosmos DB, SQL Database) and the volume of data stored.

Data transfer costs: If data needs to be transferred between services or across regions, there may be costs associated with data transfer.

Compute costs: Services such as Azure Data Factory, Azure Data bricks, Azure Synapse Analytics, and Azure HD Insight involve compute costs based on resource allocation and usage.

Security and monitoring costs: Azure Security Center and Azure Key Vault may have associated costs for advanced security features and monitoring capabilities.

Visualization tool costs: Power BI offers different pricing tiers depending on the required features and user licenses.

By considering these aspects and conducting a thorough analysis, a more accurate estimation of the project's costs can be determined.