Tabish Parkar

Summative Assessment 2: Systems Analysis And Design

05HA2309970

Bellville

In the age of digitalism, organizations constantly grip architectures that are data-driven to collect the power of various sizes of data for operational efficiency and decision making that is mindful. Moving to a data-driven system entails compound architectural decisions that will influence security, scalability and efficiency. This essay will critically analyse the key decisions which will guide overall design.

Architecture that is data-driven means that the systems that prioritize data collection, processing, storage and analysis as its main functions. This enables organizations to source information from data, decision-making management and innovation fostering. As described by sources, an architecture that is data-driven and well-designed encloses multiple elements such as acquisition of data, pipelines for processing, solutions for storage and modules for analytics which all are integrated for the supporting of organizational goals.

Scalability is vital for systems that are dealing with growing data volumes and user demands. A vital decision entails choosing the correct storage solutions. Databases such as Hadoop Distributed File System or Apache Cassandra which can scale horizontally as said by Biswas & Sen. Scaling horizontally allows for systems to grow by adding additional nodes, which verifies that storage capacity and data processing grow along with the organizational goals.

Another critical decision relates to frameworks that process data. Stream processing engines manages ingestion of real-time data along with analysis, which supports scalability with maintenance of performance. Furthermore, fostering a microservice architecture guarantees modularity, which allows for individual components to scale independently based on workload, which is vital for a multinational corporation with various data sources and user databases.

Architectural Decisions for Security

Guaranteeing that data security critical, especially for multinational corporations dealing with sensitive data along multiple jurisdictions. A vital decision entails adopting robust authentication and authorization apparatus. Role-Based Access Control and Attribute-Based Access Control policies stop any unapproved data access, which aligns with regulation standards such as HIPAA or GDPR.

Encryption also plays a significant part across levels—resting data, transitioning data and data that is currently being processed. Encrypting the data stored in dispersed databases aids in protecting towards breaches, while Secure Sockets Layer or Transport Layer Security arrangements ensure that data transmission channels are secure. Furthermore, fostering monitoring systems and audit logging enables constant facilitation, supervising early detection of security breaches.

Frameworks for data governance are also significant, founding plans for data quality, lineage and regulatory compliance. These frameworks help to guide data handling methods,

guaranteeing that protection measures are orderly imposed within the data lifecycle as described by Simmhan et al., 2018.

Architectural Decisions for Efficiency

Efficiency in systems that are data driven entails process optimization, retrieval and data storage. Partitioning data and listing contingencies critically enhance query performance, significantly in large-scale settings. For instance, partitioning data founded on time or geography minimizes query scope, minimizing resource use and improving speed.

Furthermore, fostering data compression methods lowers storage costs and hastens the transferring of data. Choosing the proper formats of data, such as ORC or Parquet also further improves efficiency by allowing for faster read or write operations.

Automation and orchestration tools like Apache Airflow or Kubernetes automate monitoring, scaling and deployment, minimizing manual interference and operational expenditure. These tools also supervise optimum utilization of resources, making sure that system performance stays efficient even during peak loads.

As the aforementioned architectural decisions adopt efficiency, security and scalability, they also present limits. For example, with distributed systems, there is an increase in complexity, which need advanced management tools and expertise. Security precautions for instance access controls and encryption can introduce latency, which effects system performance. Furthermore, growing expenditure aligned with infrastructure and constant maintenance may constrain resource-challenged organizations.

Moreover, the active behaviour of data sources and formats seeks flexible architectures that can adapt over time. Strong systems wager becoming outdated, spotlighting the criticality of modular and scalable designs which fosters future technological advancements and growth.

Designing and executing a secure, efficient and scalable data-driven system needs concise architectural decisions. Selecting the proper processing and storage frameworks helps scalability, strong security arrangements protect sensitive data and optimization methods improves efficiency of operations. While these decisions include challenges and trade-offs, a balanced strategy positioned with organizational goals and regulatory standards can increase benefits exponentially. Organizations of the future should prioritize flexible and modular architectures and constant facilitation to foster the revolutionizing data landscapes, guaranteeing a constant success in a world that is data-driven.

References

Biswas, S., & Sen, J. (2017). A Proposed Architecture for Big Data Driven Supply Chain Analytics. arXiv preprint arXiv:1705.04958. Retrieved from https://arxiv.org/pdf/1705.04958

Simmhan, Y., Ravindra, P., Chaturvedi, S., Hegde, M., & Ballamajalu, R. (2018). Towards a data-driven IoT software architecture for smart city utilities. Software: Practice and Experience. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.2580