

---

E-commerce  
Azure Cloud Architecture &  
Data Engineering Pipelines

---

Author:  
Ashish Bhavsar

## **Table of contents**

- ❖ Introduction
- ❖ Mission Statement
- ❖ Architecture Vision
- ❖ Cloud Architecture
- ❖ Pipeline Strategy
- ❖ Pipeline Failure Strategy
- ❖ Conclusion

## **Introduction**

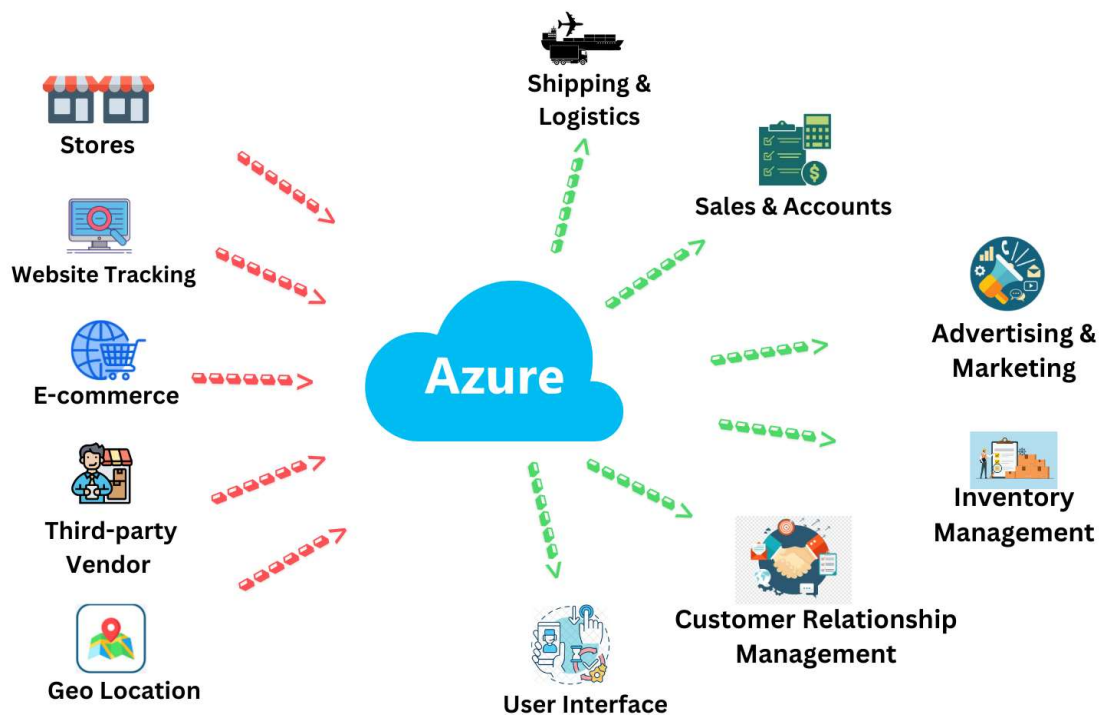
The introduction of cloud computing has caused a huge revolution in e-commerce, considering the business of Best Buy, operate online retail platforms more efficiently by implementing cloud-based solutions, which offer scalability, flexibility and cost-effectiveness. This article describes how to develop a cloud architecture plan for e-commerce with a focus on integrating Microsoft Azure to enable resilient, scalable and secure operations.

## **Mission Statement**

Our mission is to use Azure cloud technologies to build a robust and scalable e-commerce platform. This platform is intended to provide system security, high availability and adaptability to the changing needs of the market and customers. By integrating cloud resources, the platform hopes to improve overall operational efficiency and consumer happiness.

## Architecture Vision

The architecture aims to provide a cloud-native e-commerce system that connects smoothly with a variety of data sources and applications. The architecture will offer high availability, disaster recovery and scalability, guaranteeing that the platform can withstand changing loads and continue to function well in a variety of environments. The architecture will be modular, allowing for simple modifications and the inclusion of new services as the firm grows.

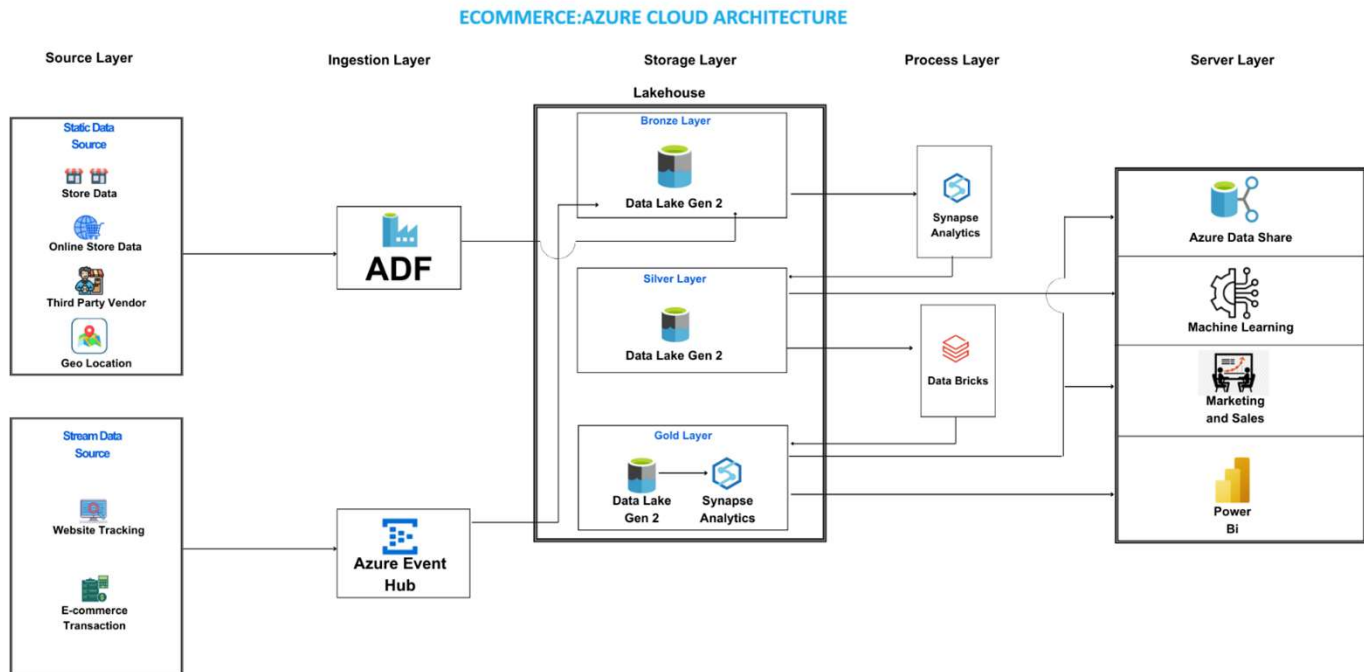


**Data Source Layer:** This layer includes many sources of data collection, such as onsite retail, online stores, website tracking, third-party vendors and geo-location services.

**Data Consumption Layer:** This layer focuses on how data is used in many corporate processes, including shipping and logistics, sales and accounting, advertising and marketing, inventory management and customer relationship management.

The architecture is modular, allowing for simple modifications and the integration of new services, and it is designed to ensure that data flows smoothly from source to point of consumption.

# Cloud Architecture



fulfilling a specialized role inside the e-commerce platform.

**1.Source layer:** This layer collects information from a variety of sources, including in-store systems, internet platforms, website analytics, third-party providers, and geographical locations.

**2.Ingestion Layer:** Data from the source layer is ingested via technologies such as Azure Data Factory and Azure Event Hub. This layer deals with both static and streaming data, ensuring that all essential information is collected in real-time or near real-time.

**3. Storage Layer:** In this tier, ingested data is safely stored in Azure Data Lake Generation 2. The data is grouped into layers, such as silver and gold, where it is cleaned, processed, and readied for future use.

**4. Process Layer:** The process layer uses Azure Synapse Analytics and Azure Databricks to convert and clean data. Data is processed and organized into selected layers, preparing it for machine learning, business intelligence, and other analytical functions.

**5. Server Layer:** The server layer manages the application and business logic. It comprises services like Azure App Services and Azure Functions, which handle the e-commerce platform's back-end operations including as database administration, payment processing, and integration with third-party services.

This multi-layered architecture ensures that data is efficiently processed, safely stored, and easily accessible for multiple business purposes, hence improving the overall performance and reliability of the e-commerce platform.



## **Pipeline Strategy**

The pipeline Strategy ensures that data flows smoothly from ingestion to final consumption.

- At first, the static data from the source layer will be ingested in 24-hour batches using Azure data factory, while the streaming data will be ingested every 5 minutes using Azure event hub.
- Since there is no immediate need for streaming data analytics, it is kept in a database rather than using Azure Streaming Analytics, which is a pricey service. Otherwise, it will be routed immediately to Azure Streaming Analytics for processing and then to PowerBI for analysis.
- Each set of data will be loaded into the first layer, the bronze layer, of the Azure data lake Gen 2, which can hold both structured and unstructured data. Datalake Gen 2 allows for easy management of both forms of data, eliminating the need for a SQL database.
- After storing, it will be transmitted to Azure Synapse Analytics for null value removal and data cleansing. The cleansed data will be saved in a silver layer in Azure Datalake Gen 2.
- The filtered data will then be delivered to Azure Databricks for further processing and aggregation. Databricks is a service used to handle large amounts of data. It is somewhat expensive.

- The fine and aggregated data are kept in the gold layer of Azure Datalake Gen2. Azure Synapse Analytics is utilized again in the gold layer to eliminate and clean up any remaining duplicates.
- The silver layer data will be delivered to the data science team for machine learning and prediction, assisting the marketing and sales teams in making informed decisions.
- The data in the gold layer will be saved for Azure Data Share, which allows for data sharing and sharing via API. In addition, utilizing PowerBI , the gold layer data will be analyzed to gain valuable insights.

## **Pipeline Failure Strategy**

A pipeline failure strategy includes constant monitoring to detect problems, creating redundancy to prevent data loss or downtime and automating recovery processes to quickly restore normal operations with minimal disturbance to business services. The strategy includes:

**Monitoring:** Pipelines are continuously monitored using Azure Monitor and Log Analytics to discover and notify on anomalies.

**Redundancy:** Using redundancy in crucial components to prevent data loss or major downtime.

**Recovery:** Automated techniques for promptly restoring regular operations in the case of a failure.

This technique ensures the pipeline's dependability and reducing the impact of any failures on company activities.

## **Conclusion**

A well-structured cloud architecture is critical to the success of an e-commerce business. The platform delivers scalability, stability and security by using Azure's cloud services, all while providing an excellent user experience. The integration of numerous data layers inside the architecture, together with a strong pipeline and failure plan, guarantees that the platform is not only capable of meeting current demands but also adapting to future difficulties. This design establishes the groundwork for long-term success in the competitive e-commerce environment.