

Cloud Architecture Of SkyMart Express

Purpose:

Design a scalable, steady, and green cloud infrastructure to help:

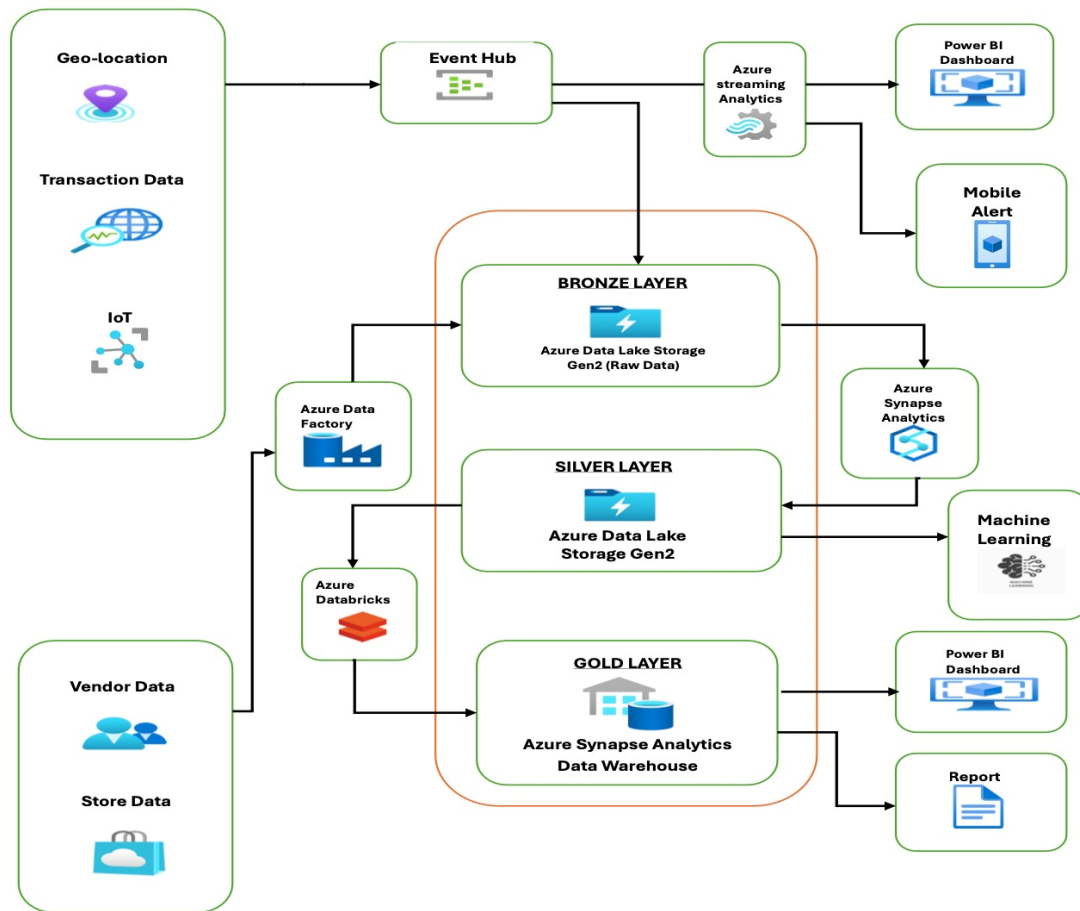
1. Online shop operations.
2. Advanced statistics evaluation desires.

Objectives:

1. Optimize Order Processing: Faster and correct transactions.
2. Real-Time Tracking: Monitor stock, orders, and purchaser actions.
3. Data-Driven Insights: Leverage trends for strategic choices.
4. Resource Utilization: Enhance the efficiency of stock and group of workers.
5. Personalization: Enable targeted advertising and person-tailored reports.

Mission

To create a reliable, secure, and scalable cloud infrastructure that supports Sky Mart's online store operations and data analytics needs.



Implementation Phases

The architecture will be implemented in three key phases, each designed to enhance data-driven decision-making and operational efficiency:

Data Collection:

- Data Sources:** In this initial phase, we will gather data from a variety of systems, including vendors, Internet of Things (IoT) devices, geo-location information, store management applications, weather APIs, and web analytics tools. This diverse range of sources will ensure comprehensive data coverage.
- Ingestion Tools:**

Azure Data Factory: This tool will facilitate both batch processing and streaming ingestion, enabling seamless data flow.

Azure Event Hub: By providing real-time updates through streaming, this tool will enhance our ability to respond to changes swiftly.

- **Storage:**
To ensure data security and accessibility, raw data will be securely stored in Azure Data Lake Storage Gen2.

Data Processing:

Data organization consists of three layers:

1. **Bronze Layer:** Represents the storage of raw data.
2. **Silver Layer:** Stores cleaned and structured data in **Azure Data Lake Storage Gen2**, ready for analytical processing.
3. **Gold Layer: Azure Synapse Analytics:** Consolidates the processed data from the **Silver Layer** and Acts as a centralized data warehouse, providing optimized data for business intelligence and reporting.

Processing Tools:

1. **Azure Databricks:** Utilized for data cleaning, transformation, and machine learning operations.
2. **Azure Synapse Analytics:** Designed for data warehousing and analytical tasks.

Visualization & Reporting:

- **Dashboards:** Power BI will be employed to create intuitive dashboards catering to various deliverables.
- **Notifications:** Mobile alerts to facilitate real-time tracking and inventory oversight.
- **Reports:** Generation of both automated and manual reports for stakeholders.

Proposed Pipeline Design: Step-by-Step Workflow

1. Data Collection:

- Ingest records from established sources (vendors/stores) and unstructured sources (IoT devices/geospatial data).

2. Raw Data Ingestion:

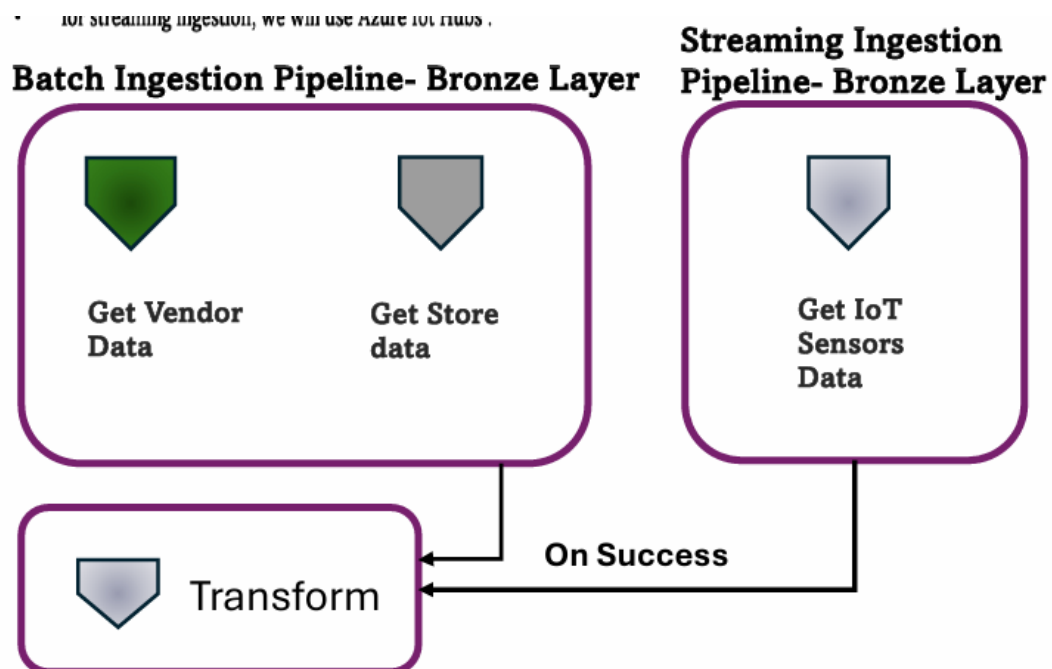
- An Event Hub captures all information streams.
- Both batch and streaming pipelines transfer data to Blob Storage or Data Lake Storage (Bronze Layer).

3. Data Transformation:

- Azure Data Factory orchestrates pipelines for:
 - Cleaning and structuring records into the Silver Layer.
 - Enriching and optimizing datasets in the Gold Layer.

4. Machine Learning Integration:

- Azure Databricks provides real-time predictions for sales trends, consumer behavior, and inventory needs.



Pipeline Components

Batch Ingestion Pipeline:

Get Vendor Data:

Retrieves established facts from seller assets, consisting of ERP structures or APIs, generally processed in batches because of periodic updates.

Get Store Data:

Collects facts from shop systems like POS (Point of Sale) or control software program, additionally in a batch mode.

Streaming Ingestion Pipeline:

Get IoT Sensors Data:

Handles actual-time streaming information from IoT gadgets, together with GPS sensors or other smart monitoring gear. This facts is ingested constantly to help stay monitoring and updates.

Transform Step

Once the information is ingested into the Bronze Layer, the Transform step cleans and organizes the raw records. This includes:

- Removing duplicates or irrelevant statistics.
- Standardizing codecs for consistency.
- Preparing the records for in addition processing inside the Silver Layer.

Control Logic: "On Success"

- If data ingestion and transformation are a success, the pipeline proceeds to the subsequent layer or subsequent strategies.
- In the event of failure, mechanisms along with retries (now not shown here but commonly part of a strong layout) can take care of mistakes earlier than moving forward.

Conclusion

This cloud structure presents a sturdy basis for Sky Mart Express to achieve operational excellence, leveraging advanced statistics processing and analytics equipment to pressure commercial enterprise selections. It is customized to guide rapid increase and adapt to dynamic marketplace demands.