Assignment 5C: Cloud Architecture

Contents

Cloud Architecture for Smart Farm: AgriTech Farms:

In this article I will outline the cloud architecture that was made for a fictious organization called ‘AgriTech Farms’. A smart farming system designed to leverage technology for improved agricultural practices. This was done to showcase how agriculture sector can take advantage of cloud architecture.

Introduction:

Agriculture is an interesting sector in regards of cloud architecture. Continuous advancement in technology have found their way into agriculture, helping it in multitude of ways. AgriTech Farms aims to integrate modern technology like sensors, drones, weather services, and cloud computing to optimize farming operations. It aims to outline a framework which will help the modern day farmers and agriculture stake holders in boosting their productivity and operations. The system provides farmers with real-time data and insights, empowering them to make informed decisions, enhance resource efficiency, and increase crop yield.

Mission and Objectives:

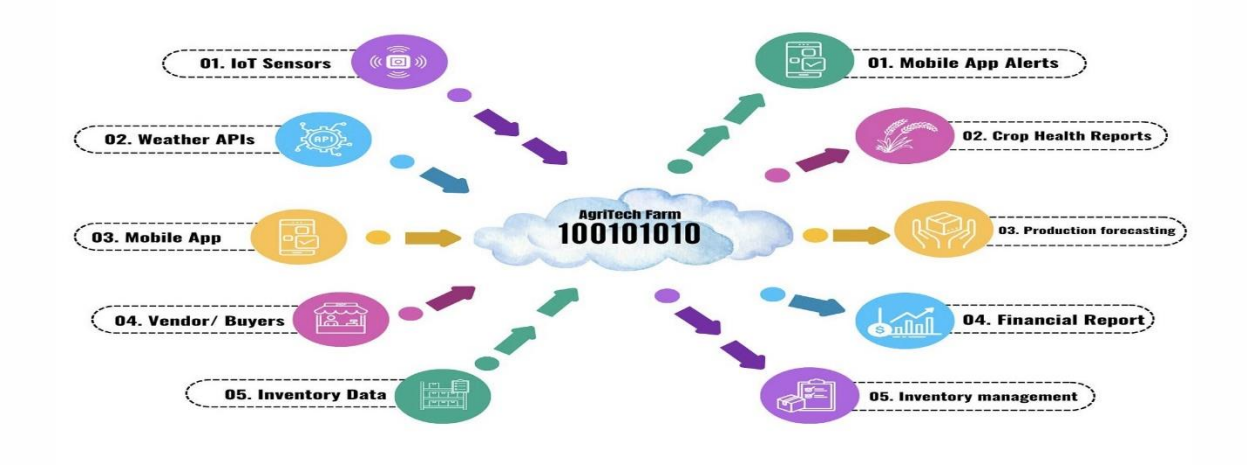
In order to help understand our scope and aims we made the following our mission which is to employ technology to transform farming into a smarter and more efficient practice. This involves using sensors, artificial intelligence (AI), and cloud systems.

The following are the key objectives, which include:

- Utilizing technology to monitor and improve farming practices.
- Conserving resources like water, fertilizer, and energy.
- Delivering real-time updates on crops to farmers.
- Mitigating crop losses and contributing to environmental protection.

Design and Discovery Phase:

For any good architecture, the design process is one of the most crucial steps. In this phase we identified the key data sources and defining the desired deliverables of the system.



We created the above diagram, to better show what would be our data sources and data sinks. The idea was to get data from farms and farmers so that we could have access to on ground data. Weather, vendors and buyers are all stakeholders or are affected by Agritech one way or another, so their data sources would also provide us with insights. And sinks were all the deliverables that will be given from the cloud.

Data Sources Overview:

For more details on the data sources, AgriTech Farms leverages diverse data sources to gather comprehensive information:

- **IoT Sensors (Soil Monitoring):** These sensors provide real-time numeric data on soil moisture, pH, and nutrient levels, streamed through an on-premises IoT Hub in a structured format. Which is essential part of maintaining healthy crops.
- **Weather APIs:** The effect of weather on agriculture cannot be understated. Hence we aim to ingest weather data from external weather APIs which will supply periodic

batch data on temperature, rainfall, and weather descriptions in structured JSON/XML format.

- **Mobile App Data:** The system's mobile app captures batch data on crop health, task updates, and user activity, syncing it to the cloud in a semi-structured JSON format. Farmers will be provided will the relevant mobile application so that they can be incentivized to collect data on the app.
- **Vendor/Buyer Data:** Information on buyers, transactions, and other vendor interactions is stored in a database and accessed in batches. This is so that we can incorporate market data or how the agriculture economy is working currently.
- **Inventory Data:** Details about stock levels and usage logs, maintained on-premises or in cloud tools, are periodically accessed in structured formats like CSV or database entries. As inventory dictates prices and can help us better allocate resources where higher financial gain can be made.

Key Deliverables:

In our cloud architecture there are a lot of key deliverables. The system aims to empower farmers with actionable information and insights through a range of features and tools, enabling them to optimize their agricultural practices and improve overall productivity.

1. Mobile App Alerts: Farmers receive real-time, actionable notifications through push alerts or SMS messages so that they can be informed of any changes or an event which may need their attention. These alerts can be regarding irrigation management so that there can be timely updates on soil moisture levels and irrigation requirements to ensure optimal water usage. Also, on pest and disease control, a proactive notifications will be sent out about potential pest infestations or disease risks based on sensor data and predictive analytics. And also, weather forecasts as instant updates on weather changes, severe weather warnings, and temperature trends will assist in decision-making.

2. Crop Health Reports: Farmers gain access to detailed insights into crop health. We plan on doing this by providing an interactive dashboard as tools like Power BI or Tableau provide live, visual overviews of crop health indicators, soil moisture, and nutrient levels. A weekly PDF Reports will contain comprehensive, easy-to-understand reports which will summarize crop performance, soil condition trends, and actionable recommendations. And for visual representation we can use heatmaps and trend analyses which will enable targeted interventions.

3. Production Forecasting: In order to help farmers plan effectively, the cloud architecture will also provide forecast dashboards which will contain predictive analytics models display anticipated yields based on current conditions, historical data, and sensor readings. It will also provide data-driven timelines optimize planting and harvesting activities for maximum yield and resource efficiency.

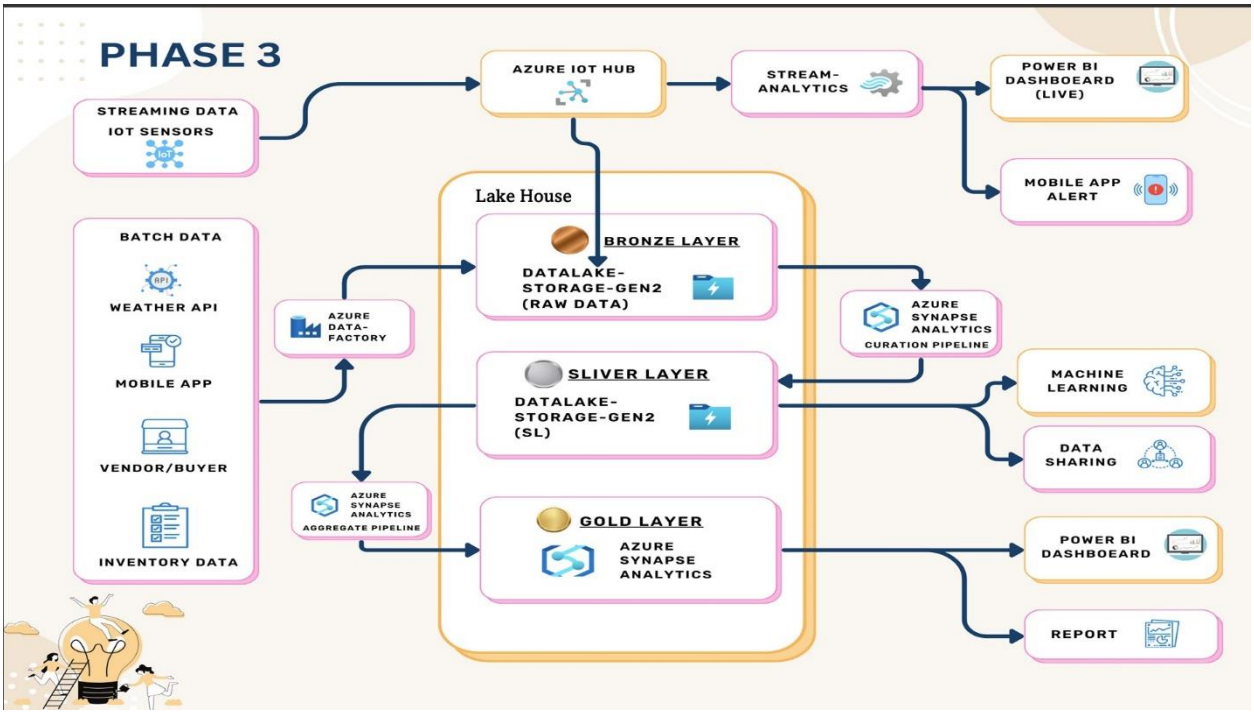
4. Financial Reports: For agronomists or any stakeholder who wants to access financial reports, the cloud will have a real-time financial dashboard which will show a detailed cost analysis, revenue tracking, and profitability insights available at a glance. A monthly email automated report will summarize profit and loss statements, key performance indicators, and expense breakdowns. And also some recommendations will be given for cost-saving measures and resource allocation strategies based on historical data trends.

5. Inventory Management: This dashboard will help in the visual representations of stock levels for seeds, fertilizers, and equipment with real-time updates. It will also provide a detailed tracking of inventory consumption patterns to predict future needs. Automated alerts will give notifications when stock levels reach critical thresholds, ensuring timely reordering.

These were the deliverables which collectively provide farmers with an integrated system for smarter decision-making, efficient resource utilization, and enhanced agricultural outcomes.

Cloud Architecture and Pipeline:

We decided to go with **Lakehouse approach** for the creation of our cloud architecture. Why Lakehouse? Because it seamlessly integrates the capabilities of data lakes and data warehouses to provide a robust, scalable solution for managing agricultural data. This approach will ensures efficient data ingestion, processing, and analysis while maintaining a clear focus on performance and reliability. To visually show our architecture the following diagram was made,



As mentioned, this was the phase 3 of our planning. The first two phases were rough drafts and after a lot of brain storming this was our final design of our cloud architecture.

To breakdown the process, the streaming data will be ingested by IoT Hub which will use DLS Gen 2 for storage and Azure Stream Analytics for transformation before posting it on the dashboard or using the data to send relevant alerts. For batch data, Azure Data Factory will ingest the data and the raw data will be stored in the bronze layer in DLS Gen 2. For more in depth understanding, the could architecture involves multiple carefully designed layers:

1. Bronze Layer (Ingest Data)

The Bronze Layer is the foundational stage of the data pipeline, this is where raw data from various sources is ingested into the system. In order to handle diverse data types and ingestion scenarios, two pipelines are employed:

- **Batch Ingestion from Azure Data Factory:**
    - Used for structured and semi-structured data sources such as historical weather datasets, financial records, and crop management files.
    - Scheduled to run at predefined intervals, ensuring timely ingestion of large datasets while maintaining system efficiency.
- **Streaming Ingestion from Azure IoT Hub:**
    - Handles high-velocity data from IoT sensors deployed in the field, capturing real-time metrics such as soil moisture, temperature, humidity, and pest activity.
    - Ensures minimal latency, enabling live data processing for immediate decision-making.

2. Silver Layer (Curated Data)

The Silver Layer focuses on data transformation and preparation, ensuring that data is accurate, consistent, and reliable for downstream analysis. The data from this layer can be used for machine learning models and if data sharing has to be done. The key processes in this layer include:

- **Data Quality Inspection.**
- **Data Cleaning.**
- **Verification and Approval.**
- **Schema Standardization.**

3. Gold Layer (Aggregate Data)

In the Gold Layer, data reaches another level of maturity and the curated data is aggregated and transformed into actionable insights. This layer is optimized for reporting and visualization, catering to specific business needs. Processes include:

- **Key Performance Indicator (KPI) Definition:** Establishing metrics to track critical aspects such as irrigation efficiency, crop health trends, and financial performance.
- **Pattern Observation:** Analyzing historical and real-time data to identify trends, correlations, and outliers.
- **Data Summarization:** Aggregating data across dimensions (e.g., time, location, crop type) to create meaningful summaries.
- **Visualization Creation:** Generating dashboards and interactive reports using tools like Power BI to present insights in an accessible and intuitive manner.

4. Pipeline Failure Handling

To ensure resilience and reliability, the architecture incorporates robust failure-handling mechanisms:

- **Retry Mechanism:** Pipelines are designed to attempt recovery from failures up to three times, with each attempt spaced one hour apart. This minimizes downtime while avoiding unnecessary strain on system resources.
    - **Event-Based Master Pipeline:** A master pipeline, scheduled to run every 24 hours, orchestrates the execution of all dependent pipelines. This pipeline is configured to handle a conditional execution based on the status of previous steps. Apart from that an automated notifications to administrators in the event of persistent failures.
- **Monitoring and Alerting:** Real-time monitoring systems track pipeline health, triggering alerts for manual intervention if required.

By employing a Lakehouse architecture with these well-defined layers and failure-handling strategies, the system ensures efficient data processing, scalability, and robustness, ultimately enabling data-driven decision-making for farmers.

Conclusion:

AgriTech Farms demonstrates how technology can revolutionize agricultural practices, promoting sustainability and efficiency. By harnessing sensors, cloud systems, and data

analytics, the system empowers farmers with real-time information and actionable insights to optimize crop production, resource management, and decision-making. The project's contribution lies in reducing waste, improving crop health, and supporting informed decision-making for a more sustainable future in agriculture.