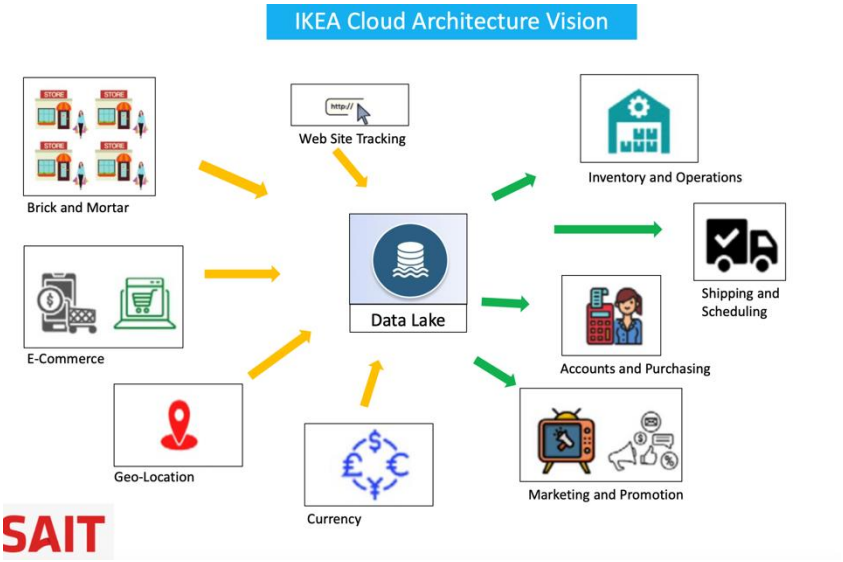
IKEA CLOUD ARCHITECHTURE

Introduction

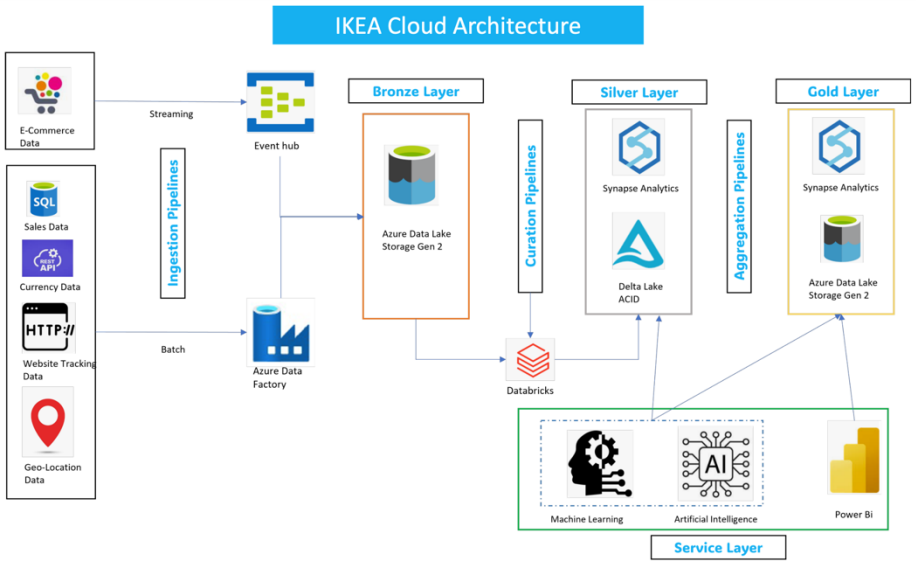
At the heart of IKEA's vision lies a centralized "data lake" where information from diverse sources such as website tracking, e-commerce transactions, geo-location, and currency data converges. This reservoir of data serves as the backbone for various critical functions including inventory management, operational logistics, shipping schedules, accounting, procurement, and marketing endeavors. By harnessing this wealth of information, IKEA aims to establish a seamlessly integrated and highly efficient system to streamline its business operations.

Vision Description



IKEA's cloud architecture envisions a central "data lake" where information flows in from various sources like website tracking, e-commerce, geo-location, and currency data. This data is then utilized for inventory management, operations, shipping schedules, accounts, purchasing, and marketing activities, creating a cohesive and efficient system for managing their business operations.

Ikea Cloud Architecture



The IKEA Cloud Architecture diagram outlines how data flows through various stages for processing and analysis. Initially, e-commerce data streams directly to an Event Hub, while sales, currency, website tracking, and geo-location data are batched through Azure Data Factory. These processes are connected via ingestion pipelines.

Both streaming and batched data then enter the Bronze Layer, stored in Azure Data Lake Storage Gen 2, before being processed in Databricks. From there, data moves through curation pipelines to the Silver Layer, which comprises Synapse Analytics and Delta Lake for transactional consistency.

Following curation, data undergoes aggregation pipelines before reaching the Gold Layer, which also consists of Synapse Analytics and Azure Data Lake Storage Gen 2 for high-performance storage. Additionally, Machine Learning and Artificial Intelligence processes are applied, connecting to both the Silver and Gold Layers for enhanced analysis.

Finally, Power BI connects directly to the Gold Layer, enabling visualization and reporting based on the refined data. This architecture ensures efficient data processing, analysis, and visualization to support IKEA's business needs.

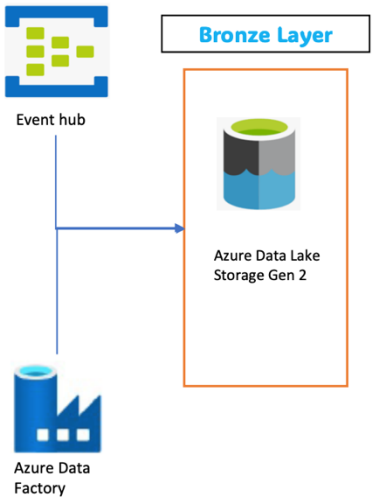
The IKEA Cloud Architecture diagram illustrates how data flows through various stages for efficient processing and analysis. Initially, e-commerce data continuously streams into an Event Hub, while sales, currency, website tracking, and geo-location data are batched through Azure Data Factory. Ingestion pipelines manage the transition between streaming and batch data, leading both streams to the Bronze Layer in Azure Data Lake Storage Gen 2, then to Data Bricks for further processing.

From there, the data moves through curation pipelines to the Silver Layer, which houses Synapse Analytics and Delta Lake Acid for refined analysis. Subsequently, aggregation pipelines further enhance the data, leading to the Gold Layer, also equipped with Synapse Analytics and Azure Data Lake Storage Gen 2, for comprehensive insights.

Additionally, there's integration with Machine Learning and Artificial Intelligence, which leverages both the Silver and Gold layers for advanced analytics and predictions. Finally, Power BI connects solely to the Gold Layer, enabling visualization and reporting based on the processed and refined data. This architecture ensures a streamlined flow of data from its raw form to valuable insights, supporting IKEA's operations and decision-making processes effectively.

Ingestion Pipelines

Ingestion pipelines in the IKEA Cloud Architecture are like the entry gates for data into the system. There are two types: batch and streaming.



For batch ingestion, data like sales figures, geographic information, currency rates, and website tracking details are collected and sent in groups, sort of like sending a bunch of emails at once. This happens through a tool called Azure Data Factory.

On the other hand, streaming ingestion is more like a constant flow of data, similar to a river that keeps flowing. Here, data from IKEA's online store is sent in real-time through something called Azure Event Hub.

Regardless of how the data comes in, everything ends up in a storage place called Azure Data Lake Storage Gen 2, where it waits to be processed further. This setup helps IKEA manage and organize all the information they gather, making it ready for the next steps in their data journey.

Curation Pipelines

Curation pipelines in the IKEA Cloud Architecture are like filters for refining data. They take the raw data stored in the Bronze Layer, where it's initially dumped, and clean it up. Think of it as tidying up messy information, making sure it's all consistent and organized.

These pipelines use a tool called Databricks, which is like a super-powered data processor, to do the cleaning and organizing work efficiently. Then, the cleaned data is stored in the Silver Layer, which is like a cleaner, more organized version of the Bronze Layer.

To ensure that this process is reliable and secure, they use Delta Lake storage, which is like a super-safe locker for data. It ensures that every step of the cleaning process is tracked and secure, so you can trust that the data you're working with is accurate and reliable.

Aggregation Pipelines

The aggregation pipelines in IKEA's cloud architecture take data from the Silver Layer, where it's already curated and processed to a certain extent. Then, using Synapse Analytics, which acts like a data warehouse, this data is further refined and combined to give a broader picture or summary of the information.

Think of it like this: imagine you have a bunch of puzzle pieces (data) that you've already sorted and cleaned up a bit (Silver Layer). Now, in the aggregation pipelines, you're putting these pieces together to create larger sections of the puzzle (aggregated data) using Synapse Analytics. This bigger picture helps you see patterns and trends more clearly.

Once this aggregated data is formed, it's stored in both Synapse Analytics and Azure Data Lake Storage Gen 2. This ensures that IKEA has access to this valuable information whenever they need it for making decisions or further analysis.

Service Layer

In IKEA's cloud architecture, the Service Layer is where specialized tasks happen. Here's a breakdown:

1. Data Marts: These are like specialized stores of data. They sit on top of the Gold Layer and directly access refined data. Think of them as focused sections in a big supermarket.

2. Machine Learning (ML) and Artificial Intelligence (AI) Teams: They work with data from both the Silver and Gold Layers. They're like chefs in a kitchen, using ingredients from different shelves to cook up insights.

3. Dashboards and Reports: They're like windows into the data world. They're created using the data from Data Marts, giving IKEA a clear view of what's happening. Think of them as fancy menus in a restaurant.

So, in simple terms, the Service Layer is where IKEA's specialists work with the best-quality data to create insights and visualizations that help make smart decisions.