TRANSORG ANALYTICS

Designing and implementing cloud data architecture and data engineering for advanced analytics at a luxury hotel chain

Introduction

A luxury hospitality company that manages a portfolio of hotels, resorts, jungle safaris, palaces, spas, and in-flight catering services wanted to migrate its customer and business data to cloud infrastructure to facilitate implementing advanced analytics use cases.

TransOrg automated the entire process of migrating data from various source systems to Azure Cloud and created a data lake framework that transforms data and creates a customer DataMart to enable advanced analytics.

Existing architecture and challenges

Client had an existing on-prem architecture with Talend for ETL processing, PostgreSQL for storage and analytics, and Power BI for reporting and visualization.

With the growth of data volume, it was cumbersome to integrate data from various sources leading to job failures and prolonging the time it takes to execute jobs resulting in delayed SLAs.
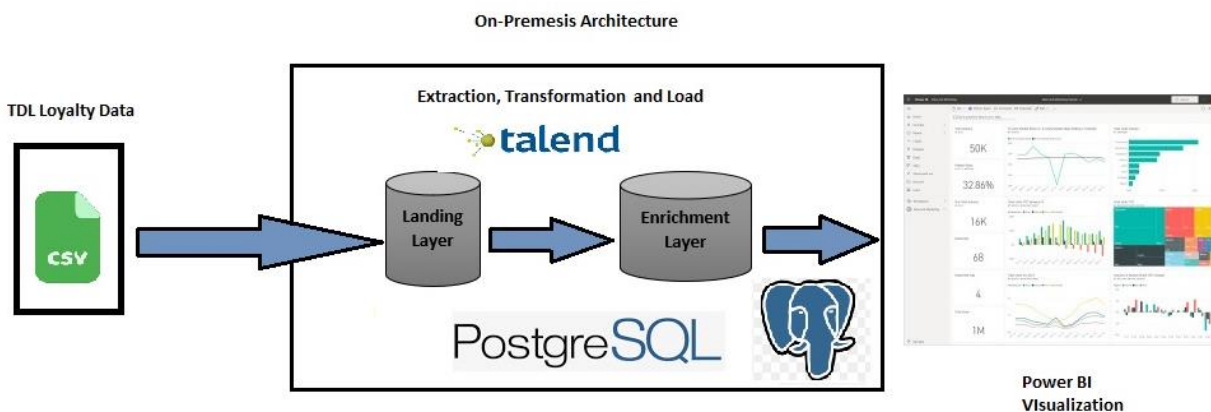

Figure 1: Illustration of client's existing infrastructure

Solution

To resolve the challenges of data volume, scalability and advanced analytics solutions, Transorg designed a data architecture that accommodates scalability of resources whenever required and can integrate more data sources to a single platform facilitating improved ETL processing and better performance.

Azure Cloud was chosen as it is one of the most accessible, scalable and on demand solutions with top notch safety and capability of integrating data from more than 60 different sources.
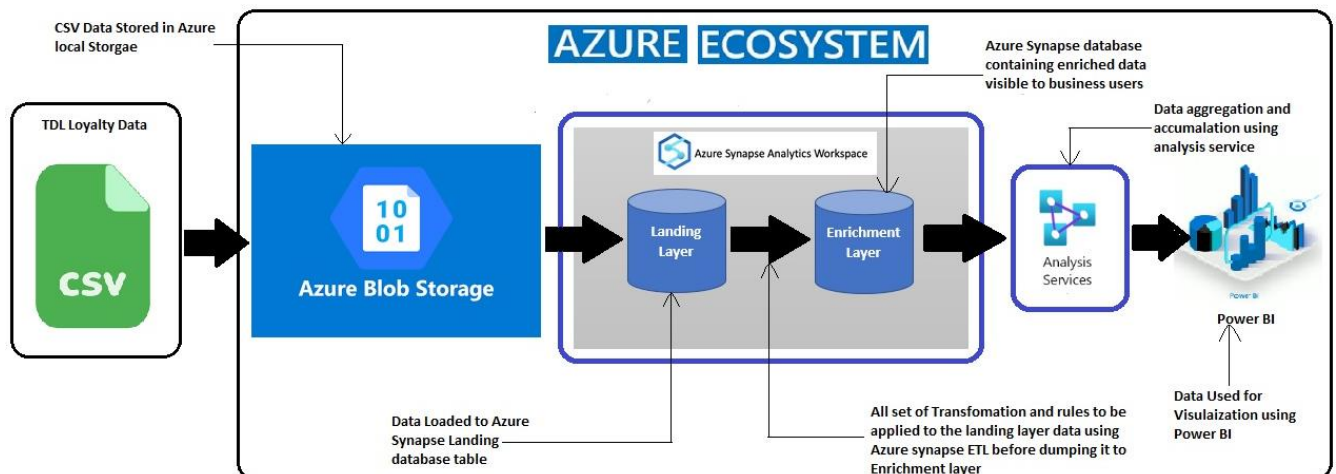

Figure 2: New infrastructure on azure cloud

Figure 2 above illustrates how the overall ingestion, processing, analysis, and visualization process of client's data on Azure platform is performed.

First, data is brought to Azure's local environment by moving files to Azure Blob Storage and stored in tables in the landing database in raw form without any modifications. Then this data is further transformed and cleansed with the incremental ETL logic implementation using Azure Synapse ETL pipelining components. Finally, data is stored in the enrichment layer where it is available for business use. Both the landing and the enrichment layer database are maintained in Azure Synapse. Data available in the enrichment layer undergoes accumulation and aggregation to populate Power BI dashboards for visualizations.

Data engineering and creation of a customer 360⁰ DataMart

Data from 14 disparate source systems is used to create a data lake framework that transforms data automatically and creates a customer DataMart. Data engineering and transformation is done using Azure Synapse during the transfer of data from the landing to the enrichment layer.

TransOrg performed the following data engineering tasks while creating the customer DataMart:

1. Data cleansing by correcting errors and replacing missing values with null values
2. Data formatting to create a uniform view of data in each column
3. Feature engineering on reservation data and deriving new columns such as financial year of customer visits, check-in and check-out month and year
4. Feature engineering on transaction data and deriving new columns such as base currency of each hotel property, latest exchange rate, total revenues, total revenues from room, food & beverages (FnB) and from other services.

After stitching, transforming, and creating the DataMart, customer 360° dashboards are created on Power BI.

Impact

Reduced duplicate customer profiles by 45%

Improved speed in data processing through Azure Cloud data lake framework

Implemented various advanced analytics use cases such as in customer segmentation, market basket analysis, personalization and demand forecasting using the DataMart

Refined decision making with insights from integrated customer 360° dashboards

Contact Us

Follow us: www.transorg.com   Email: info@transorg.com