WAVICLE
Data Solutions

Cloud Migration Brings Agility and Innovation to Cars.com, Strengthening Brand Experience

Overview

Cars.com is a leading digital marketplace and solutions provider that connects car shoppers with sellers. In today's rapidly changing market, Cars.com enables dealerships with innovative technical solutions and data-driven intelligence to better reach and influence ready-to-buy shoppers, increase inventory turn, and gain market share. On a daily basis, Cars.com hosts an average of 4 million cars from 19,000 dealers.

" Wavicle Data Solutions has been one of our most trusted partners in our cloud migration journey. They provided us with a wide variety of data engineering and AWS platform expertise, while offering significant flexibility in being able to quickly scale up our teams as the project ramped up. I highly recommend Wavicle for any company looking to embark on a cloud journey. "

– Luna Rajbhandari, Vice President, Data Management and Platforms

Customer challenge

Cars.com desired to accelerate innovation, reduce time to market, and serve highly contextual and relevant content to its shoppers and sellers. The company realized that transitioning to the cloud, rather than managing its own IT infrastructure, was critical for delivering modern business needs while reducing operational inefficiencies and overhead. The complexity of procurement and ownership, underutilization, and lead times made operating an on-premises data center costly and resource-intensive. It also limited Cars.com's ability to scale its data platform and machine learning capabilities.

Solution

With Cars.com's core business increasingly driven by data, Wavicle partnered with the Cars.com team to architect a strategic foundation for the company's entire data ecosystem. First, Wavicle helped develop an AWS Data Platform to perform data science and analytics on a wide range of data including vehicle inventory, shopper clickstream, and third-party data sources, with multiple terabytes processed weekly. Next, the existing on-premise legacy environments were re-platformed onto Amazon Web Services – which meant faster data retrieval and lower resource costs by addressing data integration inefficiencies. Re-platforming involved moving from Talend on-premises to Talend Cloud, building a data lake, and moving reporting tools such as Tableau and SAP Business Objects to the cloud. Optimizations to the architecture delivered tangible benefits when migrating to the cloud system, rather than a lift and shift approach of system migration as-is. Key to achieving this was a design decision to ingest jobs developed in Python and Spark, enabling serverless execution. As a result, job run times and Amazon EMR compute times were vastly reduced. Finally, centralizing the data in AWS Redshift will provide ease of access to data and expand the scope of reporting tools available for dealers. The data lake also integrates with a streaming platform that provides real-time insights and personalization capabilities to the Cars.com website.

Outcome

A transition of the data platform to AWS puts Cars.com in a ready position for agile innovation, time-to-market efficiencies, and service level guarantees through advanced orchestration. Matchmaking between shoppers and sellers will be enhanced. In real-time, dealers will know which vehicles are getting more attention and more likely to sell based on price and market dynamics. Additionally, new image-recognition tools built in-house will scale better on AWS in detecting vehicle attributes, such as trims and conditions, to increase the quality and accuracy of car listings.

Legacy data and technical environment

On-premise technologies included Talend, IBM Data Stage, Teradata, Hadoop, Tableau and SAP Business Objects. Source data was loaded to HDFS and Teradata, where in the processing was done by Talend and IBM Data stage. This data was eventually consumed by machine learning tools for advanced analytics, Tableau for visualization, and SAP Business Objects for adhoc reporting.

Data Sources:

- External Files (JSON, CSV, text), API, Database
- Database (Oracle, Couchbase, Salesforce)
- Streaming Data
- Cloud (Google & AWS S3)