

Case Study

Healthcare Data Nonprofit Improves Data Management with AWS

An independent nonprofit healthcare organization wanted to improve its data management strategy by migrating to AWS.



Technology

- Oracle Database
- Amazon S3 Data Storage
- Apache Parquet
- Apache Hadoop
- Amazon Elasticsearch

The Challenge

The client had already decided to migrate their analytics workloads to Amazon S3. They hoped to improve their data management strategy and, in particular, improve the user experience when accessing datasets through their portal.

The Solution

To target these issues, Datavail proposed a series of five two-week sprints working on tasks to help the client strengthen its data management posture for:

- Data ingestion
- Data transformation
- Data refresh
- Improving data security

In particular, Datavail helped the client replicate its existing Oracle on-premises database to Amazon S3 data storage—a low-cost, durable storage option.

Datavail also helped the client enact several “quality of life” improvements to enhance its ETL processes. For example, data replicated from Oracle to AWS was automatically converted to the Apache Parquet file format, so that it could be used in Hadoop large-scale data processing.

The Results

By switching to AWS, the client realized a dramatic improvement in the performance of its database queries and analysis.

- Using Amazon Elasticsearch and Kibana, users can search through 63 million rows of data in just a few seconds.
- Fully loading the database used to take multiple hours on-premises, but now takes a matter of minutes in the AWS cloud.
- Users can also now take advantage of self-service data ingestion for raw CSV files by dragging and dropping them into the ETL workflow.