



ECMC Group Excels with a Big Data Solution for Time-Series Data

INTRODUCTION

CLIENT SUCCESS PROFILE

Industry: Non-Profit/ECMC Group

Location: Minneapolis, MN

Company Size: 700 employees across the U.S.

Solution: 6 months

Technologies Used:

AWS DataLake, AWS

Redshift, AWS

QuickSight, AWS DMS,

S3, Glue

ECMC Group is a nonprofit corporation focused on helping students succeed. Headquartered in Minneapolis, ECMC Group and its family of companies provide financial tools and services, nonprofit career education and funding for innovative programs to help students achieve their academic and professional goals.

CHALLENGE

Each solution or business line at ECMC Group was originally developed independently from one another and therefore each collected and managed their own data for operational and reporting needs. Because data is so critical to the insights of a company, the CIO and his executive leadership team requested that a work team review industry best practices that would allow for cross-program analysis. Each business line was also using specialized software unique to them and there was a technical challenge with managing the entire enterprise environment. In addition, the executive team had a goal to reduce operating expense caused by the use of duplicate data management and analytical tools.

Part of the challenge also included business growth. A requirement of this growth was creating a secure and shielded environment where private company data would be isolated from external partner data.



SOLUTION

ECMC Group engaged Charter Solutions to create a platform where data from multiple sources could be centralized for reporting and analytical purposes but maintain strict access control and segregation of data. Charter Solutions designed and created an AWS Data lake architecture that allows for the ingestion of data from many different internal and external sources and segregates the data into different data domains so that access to the data domain is controlled. The business accesses their data using AWS QuickSight to pull data from an AWS Redshift data warehouse. The Data Lake architecture supports multiple layers of data to support both generalized data exploration and curated data sets to support reporting and analysis.

ECMC Group's Data Lake currently supports multiple data domains. Each data domain supports a different program or business line. Charter Solutions project team has been building the Data Lake iteratively to add data domains and incorporate additional data sources and functionality. The project team has taken the approach of creating a minimal

viable product (MVP) to meet the core business and then iteratively adding functionality to refine and expand functionality.

The initial data domain was created within 4 months of project team engagement. This domain supports the uploading of data files by 100+ external entities on a quarterly basis and brings those files through a validation process to ensure they meet minimal standards prior to being added to the Data Lake. Charter Solutions built an S3 solution to support the uploading of the files, the data file validations, and Glue processes to provide that data in a curated state for third-party reporting. Three additional data sources have been incorporated in this data domain from internal databases and external vendor APIs.

Two other domains support pulling in large quantities of data via API from third party vendors. Pulling the data into the Data Lake allows ECMC Group to create custom dashboards for the business to monitor how they are meeting both academic and operational goals as well as more in-depth analysis than was previously possible.

The most complex of the data domains involves pulling data from ECMC Group's transaction system using AWS DMS and sorting it into different data domains to support multi-tenancy. The process constantly monitors the Oracle data base for changes and pulls that into the Data Lake. The project also required the redirection of 800+ legacy Crystal reports to use AWS RedShift as the data source as well as the building of new reports and queries to support the business. The Charter Solutions team supported the reporting process by creating a solution to support the needs of the legacy reports that could not use more modern field metadata and creation of trigger processes to kick off nightly batch processing. The solution also mirrors a complex set of UAT environments used by Oracle database. This project was done in 6 months to meet the cut over deadline.

BENEFIT

The AWS Data Lake solution designed and developed by Charter Solutions in association with ECMC Group serves as a foundation to store data from disparate source systems spanning across multiple lines of business and data domains, in one place, at a low cost. As ECMC Group positions itself for future growth, the scalable AWS Analytics architecture provides the ability to add new data sources and curations in AWS, thereby eliminating redundancies, and enabling business to operate more efficiently by making better data-informed decisions.