



## CASE STUDY

# NAVIS DATA LAKE ON AWS CLOUD



## CHALLENGE

Navis data analytics team provides advanced business intelligence for hospitality customers. The current system was largely dependent on their transactional/operational database for all analytics which caused longer query times and scalability issues. Navis wanted to their customers to access rich, near-time data analytics based on democratized data which sets the foundation for machine learning for future.

## TECHNOLOGIES USED:

AWS Glue, AWS S3, AWS Fargate, AWS Athena, and Terraform.

## ABOUT NAVIS

NAVIS is the only complete customer relationship management software specifically designed for the hospitality industry. The leader in hospitality CRM solutions and services, only NAVIS provides a CRM platform that unifies your Reservations, Revenue and Marketing teams around a single source of guest data, that drives more direct bookings and more revenue.

NAVIS has been serving the hospitality business since 1987. A privately held company, NAVIS is headquartered in Bend, Oregon with growing teams in both Orlando, Florida. and Reno, Nevada.



## SOLUTION

The AWS BigData consulting team at Beyondsoft understood the desire to liberate data for analytics and proposed a data lake solution using AWS managed services and ConvergnDB (<http://convergnb.com/>), an open source solution created by Beyondsoft.

A data lake is a centralized, curated, and secured repository storing all your structured and unstructured data, at any scale. And you can run different types of analytics to better guide decision-making—from dashboards and visualizations to big data processing, real-time analytics, and machine learning.

Microsoft SQL Server is source of data with various triggers to create files as required for analytics. This process is restructured through pyspark/Glue process to store the files in S3 in bzip2 format. From S3, the data is transformed to parquet using Fargate containers running pyspark and AWS Glue ETL jobs. Using Fargate for processing files is cost efficient for smaller files as there were hundreds of small batch files to be converted. Business intelligence tool uses AWS Athena service to directly query S3 parquet files. In doing so, the records are processed to remove duplicates.

ConvergnDB is used to create AWS infrastructure using Terraform. Using a pre-built solution like ConvergnDB helps in faster time to market for the project.

The project currently involves around 0.5 tb of data from Microsoft SQL Server and converting into parquet format

for around 40 tables in 45 min batch intervals/4 times a day. Navis customers are able to access near real time data with high performance for their various reports based on the above solution.



## KNOWLEDGE TRANSFER

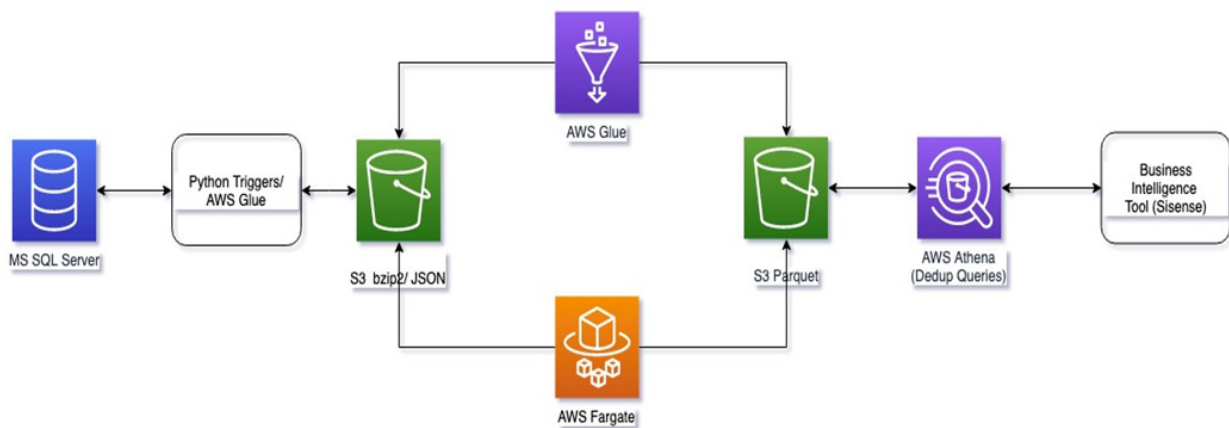
The data analytics team at Navis is educated about the AWS services, Terraform newly created solution and provided with runbook so that they can not only manage but add to the solution in future.



## BENEFITS

Navis was able to get near real time access to data so that they can take faster business decisions. Improving access to data directly translated to their customers productivity and efficiency in running their day to day operations. The project also helped to reduce the load on their transactional/operational database.

## DATA LAKE ON AWS



## ABOUT BEYONDSOFT

Beyondsoft is a global full-service IT Solutions and Service Provider. It was founded in 1995 and headquartered in Beijing, China. Beyondsoft has over 14K employees distributed over 34 nationwide offices in United States, Japan, Indian, Spain, Canada, Singapore and Malaysia. Beyondsoft is trusted AWS Advanced Tier Partner.