

Building a Serverless Data Pipeline for Jurumani with AWS



Executive Summary

Jurumani has a strong focus on Telco solutions where a recurring theme of enablement falls within the realm of Big Data. A mix of network, user, and application data builds up the portfolio of ETL and data processing requirements that they have solved for.

Challenge

Each data source has its own challenges in terms of ETL, throughput and capacity planning, storage and presentation logic. Being able to on-board a dynamic range of data sources, and managing it from a single-pane view is definitely a recurring problem in the industry. Moreover, dealing with data at this scale requires attention to the isolation of data streams and pipelines to address compliance requirements.

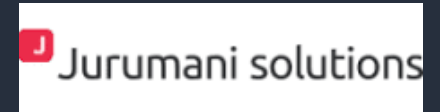
Custom-built compression and decompression logic needed to be embedded into existing Prometheus and Thanos data pipelines to accommodate a variety of consumer applications that depend on the data in various formats.

Why AWS?

AWS understands Big Data, and addresses security, scalability, service availability, performance, and cost through leading services like AWS Data Pipeline, AWS Lambda, AWS S3, and AWS Kinesis. With the level of modularity and configurability offered by AWS, each and every challenge can be perfectly countered with a managed solution with out-of-the-box automations that would have needed to be developed otherwise. AWS allows for the majority of these requirements to be solved by using Serverless technologies such as Lambda, saving massively on costs.

Jurumani chose Tangent Solutions for its strong values, and obsession with customers – a value shared by AWS. Tangent is home to a team of highly skilled technology practitioners, from cloud architects to developers and project managers, offering trust and reliability end to end. Being an Advanced Consulting Partner with a focus on Serverless Technologies, Tangent has the experience and knowledge to solve problems, while simultaneously identifying new opportunities, bringing even more value to the table.

About Jurumani



Jurumani delivers digital solutions to telecoms service providers and ISPs, designed to help engage better with customers and grow revenue.



Solution

Our solution is comprised of setting up data ingestion processes which fire off dedicated ETL pipelines with dedicated logic for each data source. Prometheus and Thanos files would be streamed into the AWS cloud by using S3 and upload commands. Serverless processing jobs would compress and decompress files depending on specific business rules. At scheduled times, CRON jobs via CloudWatch and Lambda would further compact data into other useable structures, after which analytics and presentation can be done. AWS Fargate was used to host python wrangling logic to perform additional transformations as required.

The solution Tangent provided to Jurumani is comprised of the following AWS resources:

- AWS Lambda
- AWS S3
- AWS CloudWatch
- AWS Fargate

Benefits

With the extensive use of Lambda, Fargate, and CloudWatch to automate transformation and processing jobs, a huge portion of manual intervention has been eradicated, saving up time for the Data Scientists working on the solution. Having a centralized storage system, and parallelly consumable capabilities, a culture of collaboration amongst data scientists, engineers, and managers could be established, as each would just need their own set up credentials and permissions to inspect and operate on data from their respective perspectives.

S3 and serverless has allowed not only huge cost savings (50%), but also a boost in performance, by replacing over-sized single-server monolith design practices. With parallel processing and event-driven compute, performance efficiency has resulted in data scientists spending 70% less of their time on manual intervention, error handling and reporting.

About Tangent

Tangent Solutions is a Born-in-the-Cloud Development and Cloud Consultancy company. We are an energetic company taking an innovative and disruptive approach on digital transformation, from software design to intelligent automation and Cloud. We implement technology, processes, and methodologies that make it easy for our customers to produce high quality products.

