

D&A TRANSFORMATION TALES

Decoding the power of Organized Data with **Data Lakes**



BACKGROUND

Our client is a leading flavours and fragrance company. They approached Aspire to streamline their data migration process from various sources such as files, DB, API and Stream to build a data warehouse in Snowflake within a specific time frame. The challenges they faced included:

- Corrupt, mis-formatted or incomplete data received from third-party or internal sources
- Redundant data from different divisions and company codes
- Lack of documentation support for various sources
- High data volume than expected for incremental load
- Inefficient current and history data management



BACKGROUND

To overcome these challenges our data and analytics experts suggested streamlining their data challenges with Data Lakes that would enable the client to:

- Manage a single version of truth for their reporting requirements
- Use data from Snowflake to produce BI reporting and dashboards
- Appropriate handling of historical and current data
- Scheduled and monitored job loads
- Implement data governance, data catalog and track lineage

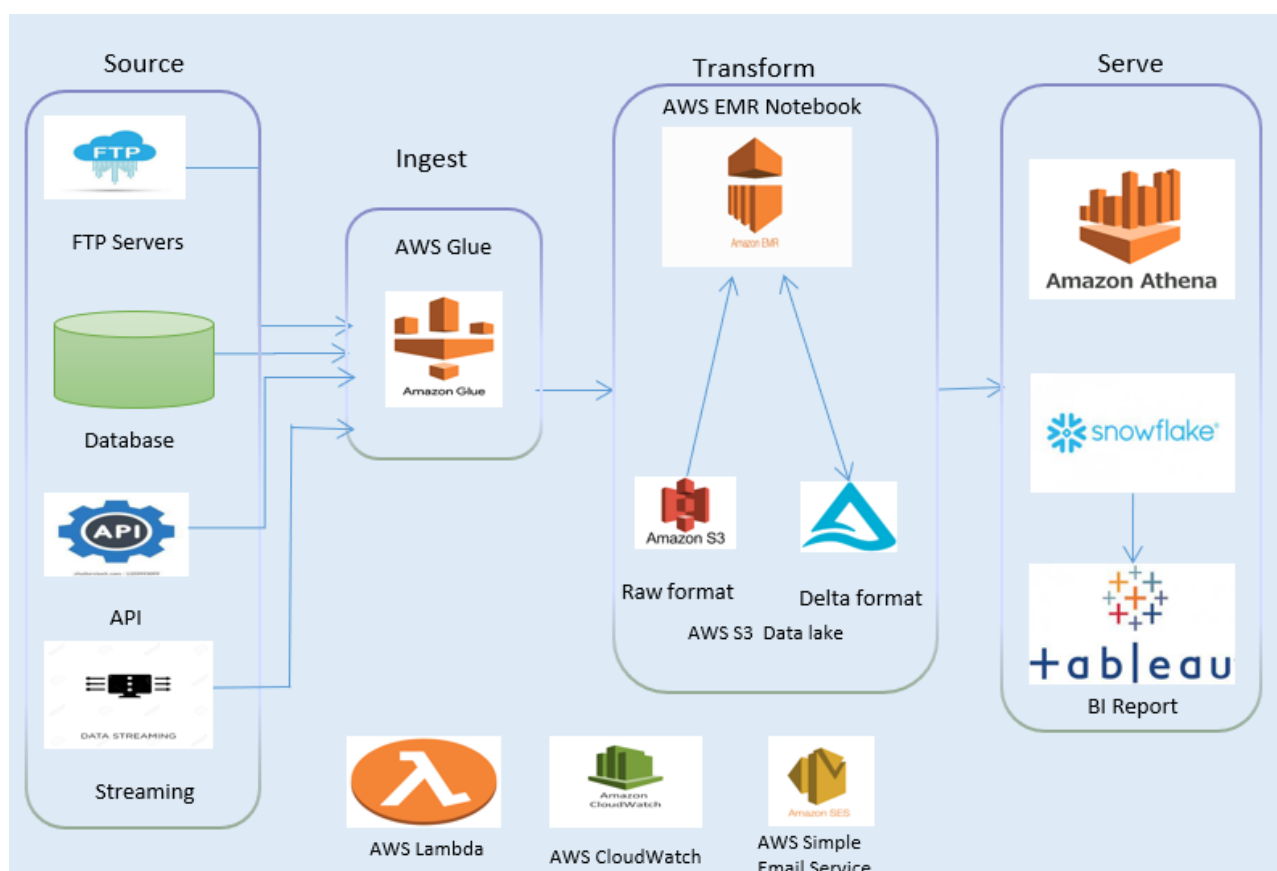




THE SOLUTION

After taking the client's needs and challenges into consideration, Aspire's Data and Analytics team curated a well-crafted plan to build and deploy Data Lake that would help the client process, store, clean, share, analyze, model and monetize their datasets using BI and ML solutions. The solution involved:

- Data Lake to leverage AWS S3 storage
- AWS EMR to implement a highly scalable and durable storage layer
- Raw zone to store ingested data in its original formats as well as data received from various sources
- Structured zone for storing data in delta lake tables, where it is populated by AWS
- EMR after the data was extracted, validated, cleaned and transformed from raw zone.
- Canonical Zone to store curated data stored in a data model, that is related data extracted from a variety of sources.
- Serving zone contains distilled data taken is directly extracted from the Canonical zone and loaded to the relevant data-stores





TECHNOLOGY SNAPSHOT

- Enterprise Cloud Platform – AWS Cloud Platform,
- AWS Glue
- AWS EMR
- AWS S3
- AWS Lambda
- AWS Cloudwatch
- AWS SES
- AWS Athena



BUSINESS BENEFITS

- Involved heterogeneous sources to produce powerful reports in milli seconds.
- High return on investment with quality reports.
- Higher query performance and improved historical insight.
- Enhanced data quality and consistency.
- Increased the power and speed of data analytics





Aspire Systems is a global technology services firm serving as a trusted technology partner for our customers. We work with some of the world's most innovative enterprises and independent software vendors, helping them leverage technology and outsourcing in our specific areas of expertise. Our core philosophy of "Attention. Always." communicates our belief in lavishing care and attention on our customer and employees.

For more info contact: info@aspresys.com or visit www.aspiresys.com

USA
+ 1 630 368 0970

SINGAPORE
+65 3163 3050

INDIA
+91 44 6740 4000

BELGIUM
+ 32 3 204 1942

NETHERLANDS
+ 31 (0)30 800 92 16

POLAND
+48 58 732 77 71

MEXICO
+52 222 980 0115