

D&A TRANSFORMATION TALES

Building an Enterprise Data Warehouse with **Advanced Reporting and Statistics**



BACKGROUND

Our client, a prominent international educational platform, faced a significant challenge in consolidating and analyzing data from its diverse and distributed system architecture. With over 40 microservice databases serving as transactional data sources, the client struggled to gain a comprehensive view of its operations and derive actionable insights for informed decision-making. Below are the major challenges our client has been looking for a promising solution for:

- **Fragmented Data Landscape:** The microservice architecture, while enabling scalability and agility, resulted in data silos across multiple databases, making it difficult to obtain a unified view of the organization's data.
- **Reporting Performance:** The transactional databases were optimized for operational workloads, leading to suboptimal performance when executing complex analytical queries and reports.
- **Data Integration Complexity:** Integrating data from numerous microservices into a centralized analytical data store posed significant challenges in terms of data mapping, transformation, and ensuring data quality and consistency.
- **Historical and Delta Load Requirements:** Capturing both historical and ongoing changes (delta loads) from the source systems was crucial for maintaining an up-to-date and comprehensive analytical data store.



THE SOLUTION

To address these challenges, the client partnered with Aspire to design and implement a scalable and performant reporting data warehouse solution. Our approach involved the following key steps:

We worked closely with the client's business stakeholders to understand their reporting requirements, analytical needs, and data consumption patterns. An enterprise-level data model was designed, adhering to industry best practices and data standards, to structure the reporting data warehouse efficiently and meet the platform's reporting requirements. After evaluating various options, we recommended leveraging open-source ETL (Extract, Transform, Load) tools to meet the client's financial constraints while ensuring scalability and flexibility.

A comprehensive mapping exercise was conducted to define the transformation rules and data lineage from the microservice source databases to the target reporting data warehouse. Leveraging Apache Airflow, a robust data orchestration tool, we developed incremental ETL pipelines to load both historical and delta data from the source systems into the reporting data warehouse.

The solution we delivered comprised the following key components:

- **Reporting Data Warehouse:** A scalable and performant data warehouse optimized for analytical workloads, consolidating data from multiple microservice databases into a centralized repository.
- **ETL Pipelines:** Efficient and reliable ETL pipelines orchestrated by Apache Airflow, facilitating the extraction, transformation, and loading of data from source systems into the reporting data warehouse.

- **Data Reconciliation and Monitoring:** Comprehensive data reconciliation processes, including data quality reports, to ensure the integrity and accuracy of the loaded data.
- **Documentation and Knowledge Transfer:** Detailed documentation and knowledge transfer sessions were conducted to empower the client's team with the necessary skills and knowledge to maintain and enhance the solution over time.





TECHNOLOGY SNAPSHOT

- Apache Airflow
- Open-source ETL
- Data Warehousing Modelling Tools



BUSINESS BENEFITS

- Unified Data View
- Improved Reporting Performance
- Scalability to accommodate growing data volumes and evolving business requirements
- Accurate and timely data at their fingertips



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@aspiresys.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