Data Platform

Dynamic data extraction and cloud adaptation with Azure Data Factory

Background

A 20-year-old software company, SI has been building and implementing SaaS technology solutions last 5 years to improve customer productivity, effectiveness, and efficiency and meet some of the most demanding contract SLAs in the world. SI is a cloud-first organization that makes extensive use of Azure Infrastructure. Headquartered in London, UK, SI offers a suite of software applications and services delivered from the cloud to provide frameworks for customers in the UK, France, and Australia. It consists of comprehensive machine-to-machine APIs as well as browser and integrated mobile application software interfaces. This SaaS software solution helps large enterprises and local governments design, build, and operate better services for their customers and users through the application.

SI relies on MSP Twisted Fish for secure database services. Having worked with many database support companies in the past, SI realized that the return on ETL processes provided by traditional approaches was low. The cost of investing in all this consulting and keeping integrations up and running was not worth the level of cloud incompatibility and inefficiency SI was experiencing in the SSIS world. Recognizing that smooth and fast ETL processes play a crucial role in internal data operations and customer satisfaction, SI knew it was critical to get data provisioning back up to speed as quickly as possible.

Challenges

In need of an advanced data extraction solution, SI experienced high resource consumption and a lack of efficiency in the progress. This led the company to reconsider the ETL tools as a machine-learned data provider. While handling the process with SSIS, SI dealt with various setbacks in terms of both cost and technique. Not only concerned about the depletion of corporate resources, SI wanted peace of mind for its customers and employees by ensuring customer higher satisfaction, reducing dependencies on infrastructure changes, and maximizing cybersecurity needs such as the ability to respond quickly to each **client's unique regulatory compliance** needs as quickly as possible.

Technology Stack

- Microsoft SSIS
- Microsoft SQL Server 2022
- Azure Data Factory
- Azure SQL Database

Key Requirements

- Incremental data flow management due to the large data.
- Automatic schema adaptation requirements.
- Quick response-ability to client's unique regulatory compliance needs.
- Ensuring security and access controls.
- Reducing dependencies on infrastructure changes.

Key Benefits

- A new ETL process with cloud-native adaptation.

- Faster and more effective data integration.

- High-quality customer experience with new Data extraction as a service (DEaS).

- Bespoke data access and data management.

- Virtually unlimited scalability for cost management.

- Agility to response client's unique regulatory compliance needs.

In addition, the company could not afford **a prolonged transition period** and the potential problems its customers might experience when adopting new tools. This meant the required new concept of ETL solutions capabilities would have to provide an advanced adoption, be easy to use, and seamlessly integrate with the existing technology stack. **The nature of the SSIS,** although powerful, is primarily on-premises tools, and the challenge is that it requires additional configuration to work with cloud environments. SSIS scalability is limited by the hardware and infrastructure of the dedicated VM resources environment.

Solution

As a part of Twisted Fish Group SI made the switch to the Data Platform's all-in-one database consultancy services once the company learned of the capabilities of the Data Platform, especially for service-level agreement.

When choosing the right concept for your ETL needs, Chris Strathie, IT Infrastructure Manager of SI, stated the three most important benefits the Data Platform provides include **scalability and performance, cost efficiency, integration, and flexibility.**

Azure Data Factory tools were used in the project to handle massive machine-learned databases and complex transformations in a cloud environment. **Client on-demand needs/requests** or app base changes will be managed with the **Managed Views** approach. Automatic adaptation requirements between source and target systems such as adding, removing, or renaming columns and tables to the database were fulfilled with **lookup tables** without **manual intervention.**

This allowed for efficient processing of big data workloads with the ability to scale up or down based on demand client requests. The nature of ADF's **parallel processing** and **watermark table** capabilities allowed faster and more efficient data movement to manage incremental data loads, and transformation by leveraging the **power of cloud computing.**

Configurational details were methodically tuned, such as **Data Integration Units** (DIUs), **throughput,** data **pipeline/activity concurrency, copy activity performance,** and **integration runtime performance.** Consequently, a proactive monitoring system was established to oversee the performance of data flows and intervene when necessary, and the service was commissioned with maximum customer satisfaction.

Key Challenges

Data was migrated to Microsoft Azure SQL Database. The migrated data was masked using dynamic data masking to comply with **GDPR.** Azure SQL Database's data masking feature obfuscated sensitive data in real-time from users, preventing data breaches and ensuring compliance. This approach allowed for the rapid development of **bespoke solutions** for each customer's unique data structure. By enhancing data security, it maintained customer trust and met regulatory requirements **cost-efficiently.**

Overall, SI reported, that the transition to Azure Data Factory was completed in only a few days with the help of Data Platform and Partner. Thanks to the Data Platform the partner's project management and their white-glove installation, training, and support at every step of the onboarding process which ensured IT personnel could get started using Azure Data Factory comfortably within one day of training. This project empowered SI to fully harness Data Factory's capabilities and achieve rapid time to value (TTV). Smooth setup helped SI quickly leverage Data Factory's powerful integration of ETL in a single solution.

- Massive and intensive machine-learned databases.

- Dynamic and complex database structure.

- Client on-demand needs/requests.

- Cloud adaptation without a prolonged transition period.

- Less flexibility and limited cloud integrations of current infrastructure.

Technical and business benefits

Data Platform exceeded SI's expectations by delivering Data Factory process and comprehensive ETL solution with high efficiency. The solution reduced the company's carbon footprint, resulting in a decreased number of devices the data personnel needed to manage. According to Chris Strathie, "Data Platform has provided many benefits such as better use of resources, the opportunity to view details on a single console and less use of capacity compared to the old solution." Data Platform simplified management of ETL and Data Factory in one place which helped Ivo and his Data team save time, bandwidth, and overhead costs. Now, SI, their Data team and employees no longer need to worry about the safety of their data and are confident the client data is overall more secure.

Data Platform

call us on: +44 7902 858497
dataplatform@rpmtechnology.co.uk

151 West Green Road, Seven Sisters,
London, England, England, N15 5EA

RPM TECHNOLOGY