



# Enterprise Data Migration - Challenges and Solutions with Snowflake

- |   |   |
|---|---|
| 1. Data migration for large retailers - challenges and importance           | 2 |
| 2. Choosing Snowflake - a future-proof solution for retailers               | 2 |
| 3. The lift & shift approach - a gentle migration of entire data warehouses | 2 |
| 4. The migration process in detail - from Teradata to Snowflake             | 3 |
| 5. Data migration leads to cost optimisation and cost transparency          | 4 |
| 6. On-time data migration to the cloud                                      | 5 |

## Introduction

Data plays a critical role in the modern business environment, especially for large retailers and enterprises that need to process a wealth of information to make informed decisions. With the growing volumes of data and the need for an agile and flexible infrastructure, organisations often face the challenge of migrating their data to a new environment. Data migration is a critical process that requires careful planning,

technical expertise and practical solutions to ensure a smooth transition. Here, we explain data migration challenges for large retailers and solutions based on Snowflake, a robust cloud data platform. Our case study describes the importance of these processes for retailers in general and the successful migration of one of Europe's largest Teradata databases to the cloud.

### 1. Data migration for large retailers - challenges and importance

In a competitive industry like retail, where customer requirements are constantly changing, and speed of decision-making is critical, a retailer's success depends heavily on the availability and

accuracy of data. However, many retailers face databases with limitations, especially regarding flexibility and scalability, and no longer meet the requirements of a data-driven world.

### 2. Choosing Snowflake - a future-proof solution for retailers

Migrating to a modern cloud data platform like Snowflake offers the opportunity to overcome these challenges and achieve several benefits. Snowflake enables flexible, scalable and future-proof data processing. It allows retailers to work with more agility and efficiency than before. In this case, it was crucial for the customer to reduce or avoid performance degradation during access peaks, i.e. inflexibility

in concurrency (simultaneous accesses). Snowflake offers additional options of interest to retailers: Security, access rights, data availability, data sharing between departments and extended functionalities. The flexible cost model, according to which the platform's operating costs are billed, also convinced our customers. With Snowflake, we hit the sweet spot between cost optimisation, scalability and flexibility.

### 3. The lift & shift approach - a gentle migration of entire data warehouses

Transferring existing data and processes to a new environment without changing the existing structures was essential for our customers. The lift & shift approach offers a gentle solution for this. We used it to migrate two independent business entities entirely to the cloud. All data structures and objects were transferred to the cloud without changing the systems. The current processes and data integrity remain intact. This is crucial for retailers, as it avoids disrupting business processes and makes the transition

seamless. While the individual departments were migrating, they seamlessly continued their work with the database. Our work was done entirely "under the hood." For example, tools that deliver data to the database have already been defined. Here, we docked to interface with Ab Initio, a data management platform, to load data into Snowflake. We connected Snowflake to MicroStrategy, a file analysis tool. This meant that data from Snowflake, rather than Teradata, was used for reporting in MicroStrategy.

## 4. The migration process in detail - from Teradata to Snowflake

### What does lift & shift mean?

In a lift & shift migration, all structures and data are migrated to the cloud without significantly changing the object structures. For the migration, this means that none of the current processes or objects were allowed to be structurally adapted.

This is particularly challenging because the architectures and processes had been optimised for the old system for years. Minimal adjustments were now allowed to build a data warehouse in Snowflake that was at least as performant.

### What steps were taken during the migration?

For the migration, the existing data and processes were divided into packages that belonged together by subject. Each package was then migrated individually. For this purpose, data structures, ETL processes, interfaces and, of course, the data were

brought into the structure required for Snowflake and then validated in terms of content. In addition, the processes and accesses were optimised considering performance. The Snowflake account was set up in parallel, including rights, warehouses, parameters, etc.

### Which tools were used for data loading & reporting?

The data was loaded with the tool Ab Initio. This was also used previously for data loading on Teradata. Accordingly, the processes triggered by Ab Initio had to be migrated to Snowflake. MicroStrategy was used as the reporting tool. The tool accesses the view layer, which was migrated to Snowflake.

An optimal setting was worked out in MicroStrategy for high-performance reporting, and warehouse architecture and settings were also developed. As a result, we were able to build a high-performance reporting system despite very large data volumes and a large number of accesses.

### What performance improvements were achieved through process optimisation and improvement of the query logic?

Due to the chosen Lift & shift migration approach and the processes optimised for Teradata, we faced the question during the migration of how to build up a performant reporting in Snowflake without significant changes to the objects. To solve the

problem, an automated materialisation framework was developed. With this, results were precalculated and cached automatically. This allowed us to achieve the required performance without making major changes to the data structures.

### How does the ACE-Deploy framework standardise the technical components and processes on Snowflake?

In the deployment framework ACE Deploy developed by initions, we have created a tool that ensures standardised deployment. With the framework, changes to the data warehouse objects occur in a controlled manner, and there is no danger to production from untested changes. For this purpose,

automated checks, approval processes, and deployment are included in the framework in predefined processes. In addition, the framework stores all database structures in a Git repository, making changes to the structures completely traceable.

### Individual authorisation concepts as best practice

For the future cloud data warehouse, an authorisation concept optimised for Snowflake was developed to control and manage all current and future accesses

easily. For this, best practices from other projects were used and adapted to the customer's circumstances.

### Test automation for performance testing and validation of a successful data migration

Since manual testing of performance and content is far from efficient, a testing framework was developed for the migration. This Python-based framework allowed us to automatically match

content and perform performance comparisons between Teradata and Snowflake. Automating the tests saved a lot of time, as only the faulty objects or slow queries had to be analysed manually.

### Cost transparency of Snowflake

In Snowflake, a warehouse architecture was built to distribute workloads as cost-efficiently as possible to individual warehouses for individual areas. This way, each area gets the computing power it needs,

and there are fewer shared resources. This makes allocating the costs incurred to the appropriate areas much more manageable. In this way, the areas only pay for what they have caused in terms of costs.

## 5. Data migration leads to cost optimisation and transparency

The successful migration was completed on schedule after 17 months with a hard deadline. Through our intensive collaboration with the customer on project planning and reporting, we created and implemented a solid migration plan. Performance targets were met. Teradata's intermittent unavailability issue was resolved through the migration. The new

Snowflake-based data warehouse creates cost transparency, e.g., between departments. There is greater scalability regarding data volumes and concurrency resilience with more users, departments or data. In the future, the database can be flexibly scaled or further developed while remaining performant.

### Establishing a new deployment framework

We established Ace Deploy, an acceleration framework developed by initions, at the customer. These technical components and processes are standardised across the entire platform at

Snowflake. With this deployment framework, clean processes for development are established. We have made the framework available to the customer for future development.

### Authorisation concepts tailored to the customer

We have tailored the authorisation management to the customer. Unique technical roles were developed for this purpose, which bundles specific rights.

This allows retailers to regulate authorisations for their data efficiently. The technical roles are used to configure data access for the departments.

### Tool for test automation

Due to unusually large data volumes and for reusability in mass testing, we built a tool to run performance tests to validate the content of the

migration reports. After the migration of each department, we performed a test to confirm that the data matched before and after the migration.

## 6. On-time data migration to the cloud

After 17 months, initions successfully migrated one of Europe's largest Teradata databases to the cloud on time. The database structure must not be changed, so a lift & shift procedure was carried out. The availability of data and access speed are improved. As Snowflake works according to the

headless principle, the new solution can be extended as our customers require and adapted to future needs. The cost model of the cloud database is based on access peaks. This enables the customer to track accesses and thus further optimise costs.



**André Paul Henkel**

Managing Director  
mindcurv: Data & AI  
Hamburg

+49 (0)40 8221 71-300