

Swinburne University of Technology
Faculty of Science, Engineering, and Technology

COS 20015 – Fundamentals of Data Management

Research report – Data Migration

PREPARED BY

Scorlib Perjuangan Lexrin

(Student ID – 102419970)

Table of Contents

| | |
|---|---|
| Abstract | 3 |
| I. Introduction | 3 |
| II. Methodology | 4 |
| 2.1 Types of Migration..... | 4 |
| 2.2 Data Migration Techniques | 5 |
| 2.3 Essential Steps to A Successful Data Migration | 5 |
| III. Discussion | 7 |
| 3.1 Why is Data Migration Difficult and Risky? | 7 |
| 3.2 Data Migration Challenges..... | 7 |
| 3.3 Data Migration vs Data Conversion vs Data Integration | 8 |
| IV. Conclusion | 8 |
| X. References..... | 9 |

Abstract

Databases, or datastores, are a crucial component in Information Communication Technologies. They allow computers to store information effectively and efficiently through scalable and logical means. One of the key consideration for a system implementation, upgrade, or consolidation is Data Migration. This report will discuss all the details such as types, planning, and techniques on how to achieve a successful Data Migration.

I. Introduction

Data migration is the process of moving data from one location to another, one format to another, or one application to another. Generally, this is the result of introducing a new system or location for the data. Data migration occurs for a variety of reasons, including server or storage equipment replacements, maintenance or upgrades, application migration, website consolidation, and data center relocation.

These days, data migrations are often started as firms move from on-premises infrastructure and applications to cloud-based storage and applications to optimize or transform a company. The following example illustrates a simple data migration flow in which the migration of Orders Delimited data of an organization takes place from a file format to SQL server following sorting with respect to Customer ID.

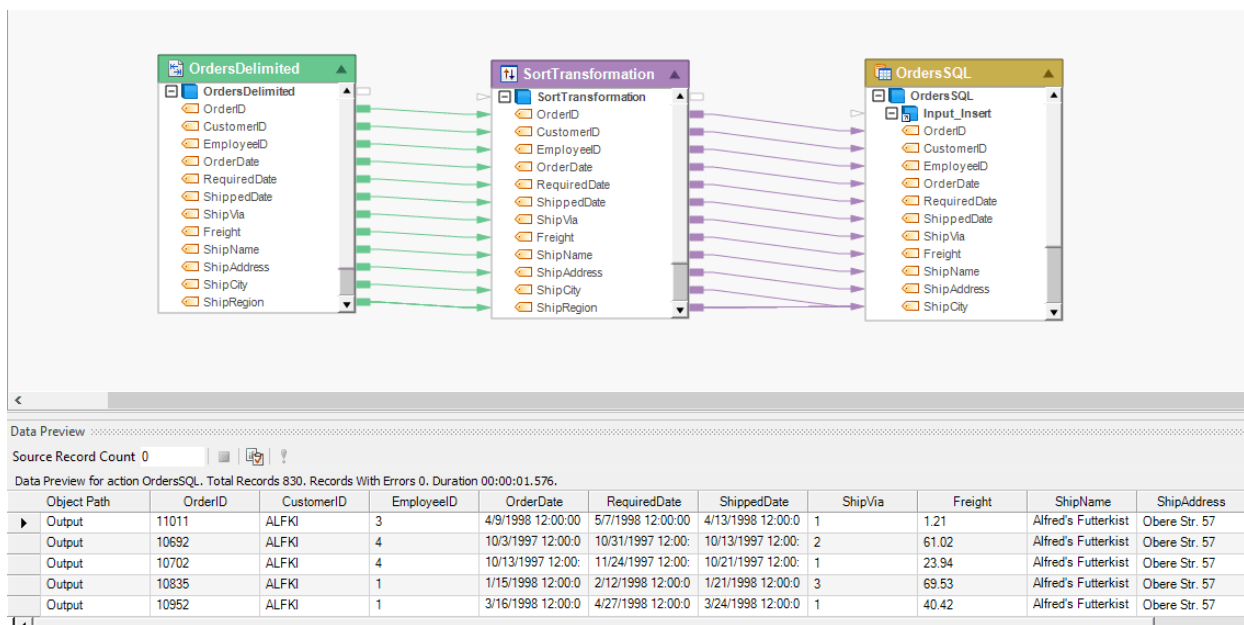


Fig1. Data Migration Example (Naeem 2020)

II. Methodology

2.1. Types of migration

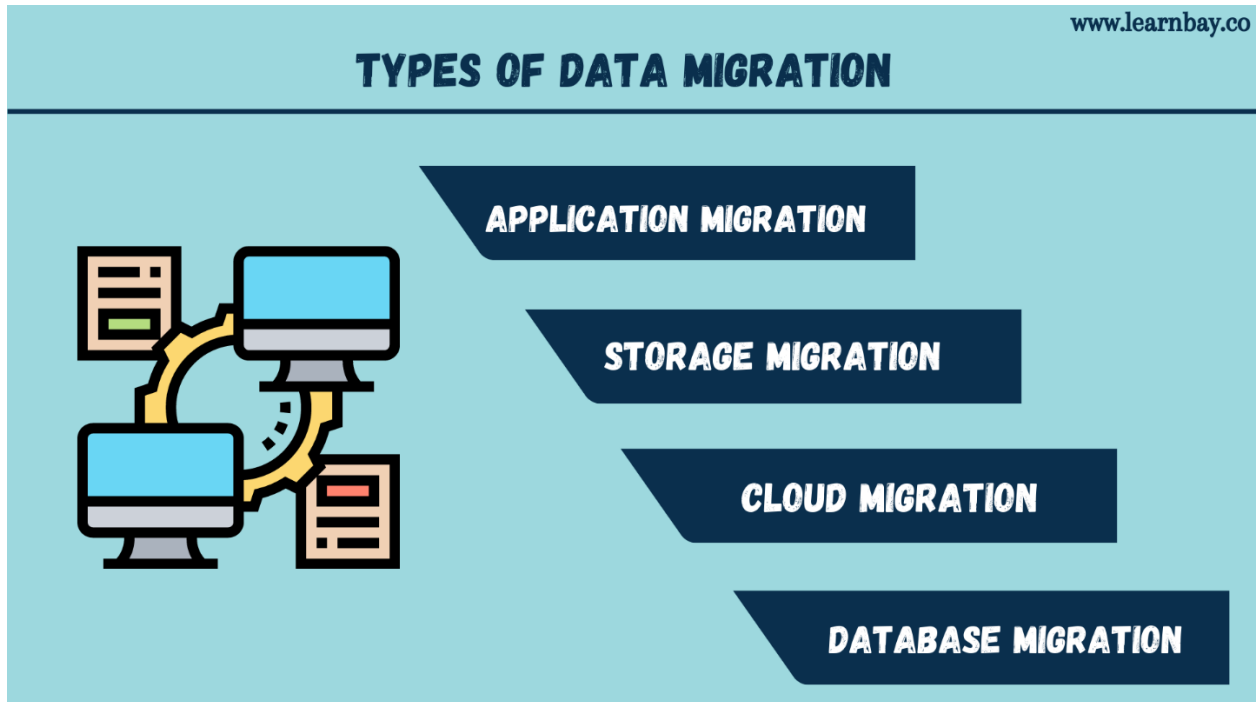


Fig 2. Types of Data Migration (Medium 2020)

- **Storage Migration.** The process of moving data from one storage system to another, such as a hard disk or the cloud. It is easy to implement data management features like data validation, cloning, reducing corrupt or old information, etc. In addition, offers significantly faster performance and more cost-effective scaling.
- **Cloud Migration.** The process of moving data, application, or other business elements from either an on-premises data center to a cloud or from one cloud to another. In cloud migration, an organization's complete or partial information assets, applications, or services are deployed to the cloud. The cloud's firewall protects the migrated data.
- **Application Migration.** The process of moving an application program from one environment to another. This occurs when an organization switches from one platform or vendor application to another. Each application has a unique data model. Moreover, applications are not portable.
- **Business Process Migration.** This relates directly to a company's business practices, often implemented by business management tools that in the case of a merger or acquisition may need to be replaced or changed. For this, movement of data can be required for anything from one store, database, or application to another.

2.2. Data Migration Techniques

- Extract, Load, Transform (ETL)

The complicated specifications of the data migration process can be managed by ETL tools. These include the analysis of large data sets, in-depth profiling of data, and multiplatform integration. The ability to automate standard ETL operations, such as extracting data from operating systems, converting it into a single format, and loading it into a destination database, is also supported by some ETL tools.

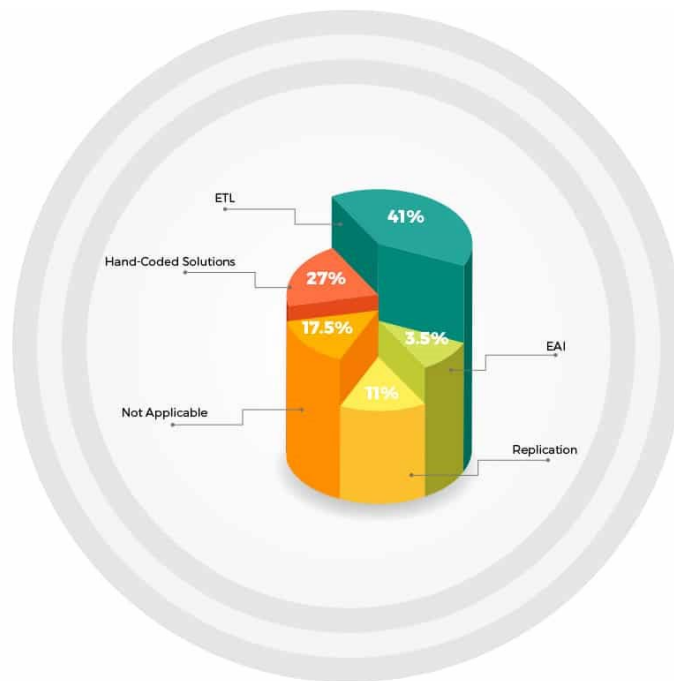


Fig 3. ETL as the Preferred Choice for Data Migration (Naeem 2020)

2.3. Essential Steps to A Successful Data Migration

Data migration is more than just moving data for one repository to another. Executing a successful transfer includes several steps. These include:

Phase 1 : Design a Strategy

- Data migration projects are complex, and a key aim is to migrate the smallest amount of data required to run the target system effectively. If a company has provided a large migration scope to the data migration project, the first stage is to define in terms of what the data sources will enable and what is fair, what is achievable. All source data is rarely needed, so scoping needs to be addressed strongly.

| |
|---|
| Phase 2 : Assess and Analyze |
| <ul style="list-style-type: none"> The next step is to analyze enterprise data. Answering the following the questions can help : <ul style="list-style-type: none"> How much busines data does migration require? Is the data unstructured or structured? Will data redundancy be a problem? How much is ROT (redundant, obsolete, or trivial) for the data? Is the information recent or old? |
| Phase 3 : Collect and Cleanse Data |
| <ul style="list-style-type: none"> This steps involves removing ROT data, compressing the remaining content, and converting enterprise-wide data into a single format. For example, data profiling transformation is applied on a data of a company to examine the available information and remove duplicate/redundant records. |
| Phase 4 : Sort Data |
| <ul style="list-style-type: none"> Once the data has been profiled into a high-quality and functional form, the next step is to categorize it according to the criteria for migration. Categorization can be performed based on the type of product, ID, or some other criterion. It makes it easier for data to be routed into the correct buckets. |
| Phase 5 : Validate Data |
| <ul style="list-style-type: none"> This phase requires evaluating the method of execution. Review the rules of the data check if they work the way they are expected to and map out any exceptions in your data flow. Even if the process is automated, it is best to keep a check on it to ensure that while implementing the process, you do not face any unexpected problems. |
| Phase 6 : Migrate |
| <ul style="list-style-type: none"> This is the final step. You will get well-organized and clean datasets from above steps; all you have to do is move them from one system to another. If the data is migrated, a database or a data archive, such as a data mart or a data warehouse, may be a target system. The employees can then use it to analyze fuel data, streamline workflows, improve data protection, and more. |

III. Discussion

This section discusses some of the theory (challenges and risks) of implementing data migration and also the comparison of with other topics such as data conversion and data integration.

3.1. Why is Data Migration difficult and risky?

The short answer is “Data Gravity”. While the idea of data gravity has been around for some time, due to data migration to cloud infrastructures, the problem is becoming more important. Data gravity is, in short, a metaphor that describes :

- How data attracts other data to it as it grows
- How data is integrated into a business company
- How data becomes customized over time

The key problem is that business processes use data in isolation and then generate their own formats, leaving the next process to be integrated. In isolation, business processes use data and then generate their own formats, leaving incorporation for the next step. Focus on the most provocative aspect of migration to ensure that the project gets the support it needs – and you will have the support of key stakeholders.

3.2. Data Migration Challenges

- **Not communicating with the business.** Be sure to keep them aware of your progress after you’ve outlined the project to the stakeholders. Every week, especially if things get off track, it’s best to provide a status update on the same day. In building trust with all those affected, daily contact goes a long way.
- **Lack of expertise.** Even though this is a straightforward process, transferring data requires a lot of difficulty. Finding an accomplished expert with outstanding references makes the process go smoothly.
- **Lack of planning.** Hours spent planning may not always guarantee results, but when it comes to actually moving the data, getting a good data migration plan saves hours. So, make sure you don’t waste your time on other unimportant things.
- **Insufficient data prep software and skills.** Invest in first-class data quality tools if this is a big migration and consider recruiting a professional company to help. Good news : To help save money, an outside company would possibly rent you the program.
- **Unproven migration technology.** Do some analysis to make sure that the procedure for data movement has performed well for other businesses like yours. Avoid the temptation of accepting only the generic procedure provided by a supplier.

3.3. Data Migration vs Data Conversion vs Data Integration

Often on the internet, many people think that Data Migration and Data Conversion are the same, they do mean different things. Data migrations is the process of transferring data between locations, formats, or systems. Data conversion in a typical data migration scenario, is only the first step in a complex process.

The term data conversion refers to the process of converting information from one format to another. This is important when transferring data. Data must be extracted from the source, modified, and loaded into the new target system based on a set of specifications in order to convert it.

Data integration is another concept that's often confused with data migration. Data integration refers to the process of integrating data from multiple sources in order to provide a single view of all data for users. Data warehouses, data lakes, and NetApp FabricPools are examples of data integration that automate data tiering between on-site data centers.

IV. Conclusion

Whatever the reason for the data migration, its ultimate purpose should be to boost corporate migration. Deliver competitive advantage and results. In order to succeed, the migration of data must be provided the attention they deserve, rather than merely being considered part of a larger underlying effort. Lacking this and without proper planning, there is a high risk that the project will go over budget, exceed allotted time, or even fail completely. The pain of handling a complex data migration can be minimized by adopting a formal approach, but to encourage a good result, the right selection of technologies can go a long way. Users should be able to implement complex business rules for data migration without requiring coding expertise or professional preparation, conversion, or data quality assurance. If you follow the steps outlined in this report, your project will have the greatest potential for success.

X. References

Learnbay Data science, 2020, 'Data Migration: Types and Popular Tools', Sep 7, viewed 17 October, <<https://medium.com/@datascience.learnbay/data-migration-types-and-popular-tools-88b1c5e66aa4>>

Oracle, 2011, 'Successful Data Migration', October, viewed 17 October, <<https://www.oracle.com/technetwork/middleware/oedq/successful-data-migration-wp-1555708.pdf>>

Naeem, 2020, 'Data Migration – The Why, The What, and The How', October 15th, viewed October 17th, <<https://www.astera.com/type/blog/data-migration-software/>>

NetApp, 2020, 'What is Data Migration', viewed October 17th, <<https://www.netapp.com/knowledge-center/what-is-data-migration/>>

Kuru, 2020, 'Data Migration', 8 July, viewed October 17th, <https://en.wikipedia.org/wiki/Data_migration>