Database Migration Strategies and Techniques to Minimize Unexpected Dysfunctionality

A Fahmi, Y H Putra

Master of Information Systems, Faculty of Post Graduate, Universitas Komputer Indonesia, Jl. Dipatiukur No. 112–116, 40132, Bandung, Jawa Barat, Indonesia

afdoli_fahmi@yahoo.com

Abstract. The purpose of this research is to develop good data migration strategies and techniques. To ensure the speed and accuracy of the process and also measure the quality of data from pre-migration and post-migration data we use the "Phased Migration Strategy" method. The results of this study are offering speed and timeliness without causing data loss or damage, maintaining consistency and integrity of the schema table structure resulting from data migration, especially in mapping data types, to match the structure in the original database, to minimize unexpected dysfunctionality. From these results in a number of other cases, a new system was developed to replace the old one which became too complicated or out of date and which held back further modification and evolution. This system is known as a “legacy” system. The rate of increase or at least a change from one system to another will not slow down in the near future. In fact, the process of migrating data from the legacy system will continue to be accelerated in the coming years.

1. Introduction

A database is a collection or collection of data that is mechanical, divided, formally defined and centrally controlled in an organization [1]. The database can be considered as a collection of files. The database system is a computerized system whose main purpose is to carry out maintenance of information and make that information available when needed. Today, databases are widely used by organizations to carry out work activities. The cost of using a DBMS (Data Base Management System) to handle a database can be an obstacle for the organization. If the organization wants to carry out its database in another DBMS, a database migration process is required. Therefore, research is needed to aim at analyzing and designing database migrations. In database migration, three main processes are needed, namely reading, converting, and writing database objects. This research helped create database replicas in the destination DBMS, although manual intervention persisted.

The process of data migration has become an ongoing problem because data has been collected. Inevitably, new systems are designed that are intended to be more efficient, user-friendly, inclusive and current (up-to-date) than older systems. In some cases, a new system was developed just to replace the older system, and the justification might be new software trends, financial constraints. New management initiatives, user complaints, or other reasons. From some examples, a new system was developed to replace the older ones that had become too complex or outdated and which resisted further change and evolution. This system is known as a "legacy" system [2]. The level of increase or at least change from one system to another will not slow down in the near future. In fact, [3] states that the process of data migration from legacy systems will continue to be accelerated in the years to come. Many organizations migrate data as part of the process of improving control, systems, or storage. Companies need to minimize the business impact of data migration such as downtime, data integrity problems, costs, control issues, and so on. The way to do this is to utilize a strong methodology for migration. The majority of data migration survey respondents reported one problem or another with this migration. Simply scheduling migration during “off-hours” is not always an adequate strategy, because:
1) The migration that goes beyond the schedule can interfere with processing during normal hours.
2) Most companies no longer have significant "off-hours" windows for activities such as data backup or data migration due to their global operations or other customer demands for availability.

There are specific objectives related to the implementation of effective data migration strategies. Especially, data must be migrated from the source platform to the target platform completely and accurately, and in accordance with company policies and regulations regarding control and information security. This means there are no records that are decreased or incomplete, and no data fields fail validation or other quality controls in the target environment. Another goal of data migration is that the process must be done quickly, with the shortest possible time. Effective data migration strategies are one-step ahead and can avoid a risk.

1.1. Definition of Database Migration
If seen from the meaning of migration, it can be interpreted as a move from one place to another, people who move from one city to another city, flocks of birds that move from cold-temperature areas to warmer regions are one illustration of what is called migration. But Migration is not only known in the context of living things, but the concept of migration is also applied to the IT field, data migration or commonly known as database migration, is one of the uses of the concept of migration used in the IT world. Data migration is a process of transferring or moving data from one environment to another, where the environment can be interpreted as database applications, technology platforms, new or different data storage locations. As an illustration, a database administrator wants to move the data he has from Oracle database applications to SQL server for a reason, the transfer process carried out by the database administrator is a process of data migration.

1.2. Reasons for Database Migration
Some basic reasons for organizations to do a database migration process are:
• The newly developed system requires an organization to use a new database application.
• Technology in the old database is too old or old.
• When viewed from the aspect of a business, it can be a long database that requires much more operational and maintenance costs than an organization must upgrade to a new system or database.
• The old system is no longer efficient to support business processes in an increasingly dynamic organization.

1.3. Challenges of Database Migration
The challenge faced by a database administrator as the executor when migrating data is how so that the old data to be migrated does not change one bit when it has been moved to a new system besides there are several other aspects that affect the successful migration of data from the old system to the system new ones include:
• Data aspect, if the complexity of the data to be moved is high both in terms of architecture and data structure, the migration process that will be carried out is even more difficult and results in the success rate of data migration
• Application Aspects, The success of data migration also depends on how different the platforms, systems, and standards of the new system, the more differences between the new system and the old system, the harder the data migration process can increase the potential for failure when migrating.
• Technology Aspects, mapping data becomes more difficult if the shape and topology of hardware and networks are increasingly spread out according to geographical or business functions.
• Human Aspects, the more human elements involved in the activity of entering, organizing, maintaining, and monitoring data, it will increase the potential for errors that affect the quality of the data stored, also increase the risk of errors that occur during the migration process.
• Policy aspects, if a company does not have clear standards as a reference in data processing, it will make it difficult for the process of determining the right migration strategy.

2. Methods
The research methodology carried out in this study is:

• "Phased Migration Strategy" to ensure the speed and accuracy of the process and also measure the quality of data from pre-migration data and post data migration.
• The TOGAF ADM Framework is only phase F, to determine or initiate a migration plan.

2.1. Database Migration Strategy
The most common database migration types are as follows:

1. Big bang Migration Strategy
   That is, all data is moved in one go and the legacy system is immediately deactivated [4].

2. Phased Migration Strategy
   Namely, data is moved in separate parts, maybe geographically or based on business functions [4].

3. Parallel Migration Strategy
   That is, data is moved, but legacy systems are allowed to run, for several periods to ensure that the target operates correctly [4].

4. Always Up Migration Strategy
   That is, migration where the source system cannot be turned off (such as a telephone system) even for a moment, this requires a special migration tool with at least advanced synchronization and complex movement from the legacy system to the target [4].

From the literature study, the comparison of data migration strategies in this study is as follows (Table 1):

| Aspect                  | Big Bang | Parallel | Phased |
|-------------------------|----------|----------|--------|
| Risk                    | High     | Low      | Medium |
| Migration Design Time   | High     | Low      | Medium |
| Execution Time Down Time| Low      | Medium   | High   |

From Table 1, it is possible to conclude that the Big Bang strategy has low implementation costs but involves higher risk, migration design time, and downtime. Conversely, Parallel strategies have a low risk, migration design time and downtime, but have high costs. For this case, the researcher chose a phased strategy because this strategy is a kind of compromise between two approaches, except that it can present a high or long migration time. However, based on the literature [10], that the problem can be avoided by making a prototype to migrate data partially in stages as shown on Figure 2.

2.2. TOGAF ADM Migration Planning (Phase F)
Migration planning is one phase of the Architectural Development Method (ADM) where migration planning will be chosen as the priority of the main work and the development of the migration plan from the old to the new system [5]. Data migration is considered the most problematic aspect of information system change, data migration can also be said to be an integral part of architectural projects. In an architectural project roadmap developed to determine the IT system to be replaced. Replacement of IT systems includes data migration activities. The relationship between the IT architecture project and data migration activities is rather complicated. The architecture framework has a direct impact on data migration activities, for example in suggesting specific shipments. More often,
this framework has an impact on IT system integration into existing architectures and indirect impacts on data migration.

3. Results and Discussion

3.1 Building a Migration Strategy

Building a migration strategy by designing data migration phase models in accordance with the evolutionary path of the application to overcome some of the problems and complexity of data migration and offer speed and accuracy during the migration process. This migration is broken down into steps to describe the migration between each situation; this step allows a more precise estimation of migration time by presenting which data and interface should be mapped between each step of the migration.

3.2 Data Migration Design

Following is the database migration design in PT. ABC environment (Figure 1):

![Data Migration Design](image)

**Figure 1.** Data Migration Design.

Figure 1 presents a data migration design that adopts from [8], an explanation of data migration design that is in the source database to move the entire object that is in the source database to the target database with the export execution mechanism, name pipe and import according to the scenario Big bang migration on the migration process.

3.3 Data Quality Review (DQR)

In accordance with the methodology described, the first model of the data migration process is a review of data quality using DQR (Data Quality Review) adopting the practical data migration [4], Table 2 is DQR form, standard from the practical data migration [4], that has been completed in accordance with the research environment, namely PT. ABC is as follows:
Table 2. DQR Form.

| Identity |
|----------|
| Short Name | ASA001 |
| Raised by | Afdoli Fahmi |
| Entity | Divisi Sistem Informasi |
| Date raised | 01/01/2018 |
| DQR ID | 1001 |
| Priority | High |
| Version | V.1001 |
| Status | In Progress |

Data Quality Assessment

Qualitative assessment
From the qualitative assessment stage, checking data migration is based on data consistency and redundancy of data in the source database.

Quantitative assessment
From the quantitative assessment stage, checking the migration of data is based on the number of tables up to the number of records in each table carried out before the process of migrating data from the source database to the target database.

Method

| DQR tasks | Description | Who | When |
|-----------|-------------|-----|------|
| Verify | Verification of data by checking the number of tables and records at source | Afdoli Fahmi / Sisfo Division | Monday, March 5, 2018, until Wednesday, March 7, 2018 |
| Retention | Retention checks from the age of the data on the source of | Tri A Gunawan / Membership Division | Monday, March 5, 2018, until Friday, March 9, 2018 |
| consistency and redundancy | Check Consistency and data redundancy at source | Afdoli Fahmi / Sisfo Division | Thursday, March 8, 2018, until Saturday, March 10, 2018 |

Method statement
DQR Task is done before the process of migrating data from the source database to the target database, after executing data migration, the DQR task is done again to ensure the quality of the data from the migration process in the target database.

3.4. Data Migration Executed
At the stage of data migration development. Sequential migration mechanism namely the move runs sequentially or gradually were carried out preceded in the pension payment transaction table. this sorting is sorted per paid partner available in the transaction table with data that has a small number of
records first then the number of records medium and last with more number of records, this is done to test the success of the migrated data with the processing time taken can be measured according to the validation carried out.

### 3.5. Migration Planning

The Migration planning phase at TOGAF has the purpose to plan the migration process or transition from the current state or the source database server to the proposal or target database server with a design based on enterprise architecture so that the application becomes directed and runs well. This migration process is through an implementation roadmap plan.

#### 3.5.1 IT Design Architecture to Data Migration

The TOGAF Framework has a direct impact on IT system integration into existing architectures and indirect impacts on data migration; in the case of this study the authors describe the roadmap for implementing data migration as follows (Figure 3):

![Figure 2. From IT Architecture to Data Migration.](image)

Figure 3 Data migration is related to enterprise architecture development because there is a direct or indirect connection when the architecture or system accesses the database.

### 3.6. Decommission Legacy

The final step is to turn off legacy or old systems and realize data migration, which ends with the cutting of legacy systems. Usually there will be a number of prerequisites that need to be met before the old system can be stopped. Ensure that it is fully documented and agreed upon (should be done in the previous phase) so that it can begin to validate that migration has met the requirements. Thus, data conversion experts must remember the following problems:

- Freeze of legacy system for future use.
- A final copy is prepared.
- The old system provided in "view" mode.
- Archiving historical data.

### 4. Conclusion

The objectives of this study are to develop strategies and techniques of good and correct data migration that offer speed and timeliness without causing loss or damage to data, maintaining consistency and integrity of the schema table structure resulting from data migration, especially in mapping data types, to fit with structures in the original database, to minimize unexpected dysfunctionality.
Acknowledgments
Author would like to thank the support from DR. Ir. Yeffry Handoko Putra, M.T as the Head of Magister of Information System Department UNIKOM who provided advice and insights during this research and to all parties from Universitas Komputer Indonesia that involved in writing this report either directly or indirectly.

References
[1] Gordon C. Everest. (1986). Database Management: Objectives, System Function and Administration : McGraw-Hill.
[2] Brodie M. and Stonebraker M. (1995). “Migrating Legacy Systems Gateways, Interfaces and the Incremental Approach”: Morgan Kaufmann.
[3] Breeding, M. and Roddy, C. (2003), “The competition heats up” Library Journal, April 1, pp. 52-64
[4] John Morris. (2009). Practical Data Migration: Second Edition
[5] A Josey and the TO Group 2011 TOGAF ® Version 9.1 Enterprise Edition - An Introduction (San Fransisco CA).
[6] Open Group. (2009). The Open Group Architecture Framework: Architecture Development Method. http://www.opengroup.org/architecture/togaf9-doc/arch/.
[7] Lussem, J., and Harrach, H. (2013). How to make data migration processes more efficient by using TOGAF: Best practice data migration approach applied to SAP Financial Services-Policy Management. 2013 ACS International Conference on Computer Systems and Applications (AICCSA).
[8] F. Matthes, K. Haller and C. Schulz, "Testing and quality assurance in data migration projects," 2011 27th IEEE International Conference on Software Maintenance (ICSM), Williamsburg, VA, USA, 2011, pp. 438-447.
[9] Cala, A., Luder, A., Cachada, A., Pires, F., Barbosa, J., Leitao, P., and Gepp, M. (2017). Migration from traditional towards cyber-physical production systems. 2017 IEEE 15th International Conference on Industrial Informatics (INDIN).
[10] Erik, P., Bansleben. (2004). Database Migration: A Literature Review and Case Study.