

Advanced Migration of Schema and Data across Multiple Databases

D.M.W.E. Dissanayake

139163B

Faculty of Information Technology

University of Moratuwa

May 2017

Advanced Migration of Schema and Data across Multiple Databases

D.M.W.E. Dissanayake

139163B

Thesis submitted to the Faculty of Information Technology,
University of Moratuwa, Sri Lanka for the partial fulfillment of the
requirements of the Degree of Master of Science in
Information Technology.

May 2017

Declaration

We declare that this thesis is our own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Name of Student

Signature of Student

Date:

Supervised by

Name of Supervisor

Signature of Supervisor

Date:

Dedication

I dedicate this thesis to my dear parents for being great role models and encouraging me every day to achieve my goals in life.

Acknowledgements

I would like to thank my thesis supervisor Mr. S. C. Premaratne of the Faculty of Information Technology at University of Moratuwa, for the guidance and support he provided throughout the whole research.

Abstract

Migration of existing databases is an unavoidable task in modern technical and business environment. Most of the time, the DBAs are forced to work with a tight budget to choose the ideal software to accomplish the migration requirement of their organization and complete the process within a limited time period.

When considering the commercial and freely available data migration tools we found a number of disadvantages. Some of the tools are very expensive. Therefore the management of small scale companies may not approve purchasing such software unless they are convinced it's a necessary expenditure as well as a good investment for the future. Even though they offer many features, some of the tools are suitable only for cloud based migrations. Some of the tools can only perform the migration process, inwards. Some tools are not available as standalone programs, so the customer has to purchase the whole software package which the required tool is included. Some tools don't provide GUI for the users. Therefore the user must have good programming knowledge in order to work with it.

In order to overcome these issues, I. M. Wijewardana presented a solution which offers an open and extensible migration process with attractive and user-friendly interfaces. However there are some important features yet to be implemented, including the features which have been suggested as future work by Wijewardana and colleagues.

With this research, we have extended the work of I.M. Wijewardana and implemented the following features. Handling the effects of primary key change of a table in between two data migration sessions. Informing the relevant details to the user or take action based on user preferences in an event of existing data replacement. Incremental update of data: Facility to update tables column-wise as well as row-wise, without causing any loss of data. Handling changes when the number of columns in a table (in source database / target database) gets increased / decreased. Handling changes when the constraints of table columns get changed. Facility to migrate the parent tables of a table which were initially selected to migrate based on relationships between the tables. Handling table and column name changes, column mapping, data type mapping. Migrating data based on different criteria (Data selection via SQL queries). Rollback database to a stable / pre-migration state in an event of unsuccessful data migration

Contents

	Page
Chapter 1 – Introduction	01
1.1 Prolegomena	01
1.2 Background and Motivation	02
1.3 Problem Definition	02
1.4 Aim and Objectives	02
1.5 Structure of the Thesis	03
1.6 Summary	03
Chapter 2 - Current Development in Database Migration	04
2.1 Introduction	04
2.2 A Methodology for Data Migration Between Different DBMS	05
2.3 A Framework for Data Migration Between Various Types of RDBMS	06
2.4 Transform! Patterns for Data Migration	08
2.5 A Metadata Driven Approach to Performing Complex DB Schema migrations	09
2.6 Quickie: Automatic Schema Matching for Data Migration Projects	10
2.7 Criteria for Evaluating General Database Migration Tools	12
2.8 Migrate and Transfer Schema and Data across Multiple Databases	14
2.9 Problem definition - Research Question	15
2.10 Summary	15
Chapter 3 - Technology Adopted for Database Migration	16
3.1 Introduction	16
3.2 NetBeans Platform	16
3.3 MySQL Workbench	18
3.4 pgAdmin	19
3.5 SQL Server Management Studio	20
3.6 Oracle SQL Developer	21
3.7 The Java Database Connectivity (JDBC)	22
3.8 Summary	22

Chapter 4 - A Novel Approach to Database Migration	23
4.1 Introduction	23
4.2 Features of the previous approach	23
4.3 Features of the new approach	23
4.4 Inputs	24
4.5 Outputs	25
4.6 Process	26
4.6.1 Schema and Data Migration Process	26
4.6.2 Incremental Data Updates: Column-wise	27
4.6.3 Incremental Data Updates: Row-wise	27
4.7 Assumptions	28
4.8 Users	28
4.9 Summary	28
Chapter 5 - Solution Design	29
5.1 Introduction	29
5.2 Interaction Between Objects	29
5.2.1 Interaction Between Objects: Selecting source database	30
5.2.2 Interaction Between Objects: Selecting tables	31
5.2.3 Interaction Between Objects: Mapping columns of source and target tables	32
5.2.4 Interaction Between Objects: Resolving issues	33
5.2.5 Interaction Between Objects: Managing data migration criteria	34
5.2.6 Interaction Between Objects: Executing updates	35
5.2.7 Interaction Between Objects: Reviewing migration progress	36
5.3 Database Connectivity Architecture	37
5.4 Summary	38
Chapter 6 - Implementation of the solution	39
6.1 Introduction	39
6.2 Interface DatabaseMetaData	39
6.3 Interface ResultSetMetaData	39
6.4 Implementation	40
6.5 Summary	48

Chapter 7 - Product Evaluation	49
7.1 Introduction	49
7.2 Advantages of the features in our new approach	49
7.3 Comparison of features available in old approach and new approach	51
7.4 Actions taken by the system in various difficult scenarios	52
7.5 Summary	55
Chapter 8 - Conclusion	56
8.1 Introduction	56
8.2 Results	56
8.3 Further Work	60
8.4 Summary	60
Chapter 9 - References	61

List of Figures

Figure 3.1: MySQL Workbench main screen window	18
Figure 3.2: pgAdmin III for PostgreSQL	19
Figure 3.3: Microsoft SQL Server Management Studio	20
Figure 3.4: SQL Developer main window	21
Figure 4.1: Methodology of Advanced Schema and Data Migration System	25
Figure 5.1: Sequence Diagram - Selecting source database	30
Figure 5.2: Sequence Diagram - Selecting tables	31
Figure 5.3: Sequence Diagram - Mapping columns of source and target tables	32
Figure 5.4: Sequence Diagram - Resolving issues	33
Figure 5.5: Sequence Diagram - Managing data migration criteria	34
Figure 5.6: Sequence Diagram - Executing updates	35
Figure 5.7: Sequence Diagram - Reviewing migration progress	36
Figure 5.8: Database Connectivity Diagram	37

Figure 6.1: Data type mapping	40
Figure 6.2: Selecting source database system	41
Figure 6.3: Selecting a particular database in source database system	41
Figure 6.4: Selecting target database system	42
Figure 6.5: Selecting a particular database in target database system	42
Figure 6.6: Selecting tables	43
Figure 6.7: Mapping columns of source and target tables	43
Figure 6.8: Resolving issues: Part 1	44
Figure 6.9: Resolving issues: Part 2	45
Figure 6.10: Managing data migration criteria: Part 1	46
Figure 6.11: Managing data migration criteria: Part 2	46
Figure 6.12: Executing updates	47
Figure 6.13: Reviewing migration progress	47
Figure 8.1: Results: Oracle Database	56
Figure 8.2: Results: MySQL database	57
Figure 8.3: Results: MS SQL Server database	58
Figure 8.4: Results: PostgreSQL database	59

List of Tables

Table 7.1: Comparison of features available in old approach and new approach	51
--	----