SECTION II

ENGINEER - Vol. XXXXV, No. 02, pp. [41-48], 2012

Evaluation and Selection of Tools for Data Migration from Non-Spatial to Spatially Referenced Software – A Case Study Migration from MySQL to PostgreSQL

E. A. G. Chandramali and N. T. S. Wijesekera

Abstract: Geographic Information Systems and open source software are becoming more and more popular resulting in an increased requirement for data migration from common non spatial software to spatially referenced software. MySQL is a very popular open source Database Management System used by most web developers but without support for spatial referencing. PostgreSQL is an open source software that supports Geographic Information Systems. It is often necessary to migrate data from MySQL to PostgreSQL and it is possible to identify many tools that are capable of executing the desired task. Therefore best available tool should be selected to ensure that the selected tool satisfies the main functionalities expected of the software and capabilities of performing the tasks with user friendly features. This paper describes a systematic methodology adopted to select the best free tool for data migration from MySQL to PostgreSQL using literature and rational judgement incorporating a qualitative ranking system to Migrate the base data. For easy comparison, an Evaluation Score of Tool was defined by calculating the percentage of available functionality when compared with the user desires for satisfaction. After comparing three available tools, Postgres Plus 8.3 free software was identified as the best with a Evaluation Score of Tool value of 76%. The present work identified 3 main and 11 sub database functions together with 4 main and 13 sub components of Graphical User Interface functionality as important parameters for data migration.

Keywords: MySQL, PostgreSQL, Open source, Data Migration,-Tool Evaluation, Web Mapping, GIS

1. Introduction

Spatially referenced map based information systems and open source software are becoming more and more popular resulting in an increased requirement for data migration from already developed common non-spatial software to spatially referenced software.

MySQL is a Relational Database Management System (DBMS) that runs as a server providing multi-user access to a number of databases [1]. PostgreSQL which is often called Postgres, is an object-relational database management system and is free and open source software [2]. PostgreSQL is increasingly used for large open source business applications. Due to the speed of data operations and the availability of features that work well with web-based servers, MySQL is preferred by most web developers for the use of database management software [4][5][6].

In case of map based web applications, developers prefer the use of PostgreSQL as it supports the use of spatial referencing which is a key in the development of Geographical Information System (GIS) applications. As an example, MySQL has very limited GIS capability such as with Minimum Bounding Rectangle (MBR), while PostgreSQL provides support with full geospatial capability (PostGIS) that conforms to the OpenGIS standards. While MySQL is commonly known as the most popular open source web database, PostgreSQL is referred to as the world's most advanced open source database or the open source Oracle [6]. As such for GIS based web applications, it is necessary to convert non spatial data developed with MySQL to PstgreSQL and this has to be done easily and effectively.

There are many reasons for the use of PostgreSQL over MySQL. PostgreSQL supports complex database design with its advanced rule sets, enables use of procedures

Ms. E.A.G.Chandramali *, B.Sc. (Hons)in IT . (SLIIT), M.Sc (Moratuwa), Research Assistant, Department of Civil Engineering, University of Moratuwa, Sri Lanka*

Eng. (Prof.) N.T.S.Wijesekera*, B.Sc. Eng. (Sri Lanka), M.Eng. (Tokyo), D. Eng. (Tokyo), C. Eng., FIE(Sri Lanka), Senior Professor of Civil Engineering, Department of Civil Engineering, University of Moratuwa, Sri Lanka*

to be executed by the database server, has many transactions, permits the use of stored procedures, does support R-Trees indexes which can be used to improve the efficiency of the nested sets, are some of them [7].

There are many ways that can be used for migration of data from MySQL to PostgreSQL. However a GIS inclined user does not appear to have a suitable guidance as to how such a migration could be carried out ensuring the data accuracy, spending minimum time, and avoiding frustration.

Migration from MySQL to PostgreSQL for the need to incorporate spatial referencing in a GIS would either mean migrating both the application and database or capturing the basedata and then redesigning the application. Difficulties of MySQL to PostgreSQL are cited by many and it is stated that a complete migration, for an example, when migrating for performance reasons, the developer has to change many other components. In such situations, it would require at least 3 months [8].

A literature survey captured a set of tools as shown in the Table 1

Table 1 - Identified Tools for Data Migration from MySQL to PostgreSQL

Item	Name Given for the Tool	Free/ Proprietary*/ Open Source	Source
1	SwisSQl- Data Migration Tool 6.0	Proprietary ---- (30 day trial is available)	swisSQL http://www.swissql.com/products/datamigration/data-migration.html
2	Data Loader ver 4.0	Proprietary – limited functions – allowed to load up to 50 rows	Data Loader http://www.dbload.com/
3	DB Convert	Proprietary – trail version is available	Db convert http://dbconvert.com/convert-mysql-to-postgresql-pro.php
4	Migrate Data 1.1	Free and Open source	Migrate Data 1.1 , Softpedia http://www.softpedia.com/get/Programming/Other-Programming-Files/Migrate-Data.shtml
5	pgEDIT	Proprietary Available for free download - Free only to try	Pgedit http://pgedit.com/download
6	Postgresplus 8.3	Free and Open Source	EnterpriceDB http://www.enterprisedb.com/products /download.do
7	ESF Database Converter	Free and Open source	ESF Databse Converter http://www.easyfrom.net/download/
8	MySQL2PostgreSQL, PRO 1.3.0	Proprietary	Dbconvert http://3d2f.com/programs/49-269-mysql2postgresql-pro-download.shtml
9	MySQL PostgreSQL Import, Export & Convert Software 7.0	Proprietary	Sobolsoft http://3d2f.com/programs/58-642-mysql-postgresql-import-export-convert-software-download.shtml

** Licenced software available in the market for a price*

Present study was carried out to identify the best method to evaluate the available free tool options ensuring the user needs such as the efficiency, accuracy and the user friendliness and then to evaluate the captured set of tools for data migration from MySQL to PostgreSQL.

A brief description of the selected tools are as given below. Postgres Plus 8.3 is an open source, pre-configured, certified binary distribution that simplifies enterprise deployment, eliminating the need to manually assemble and integrate software components from a variety of web locations. This supports most popular environments, including Linux 32 and 64, Windows, and Mac. It is said that data migration from MySQL can be easily handled with this tool. Migrate-Data 1.1 is a pure java tool and an extremely powerful enterprise data migration tool aimed at small and medium sized databases. It is indicated that with Migrate-Data a user can easily extract, transform, load and integrate data from any-to-any (any data source into any database) databases. In addition, it is stated that the tool enables a user to perform customized data extraction from source and create customized target objects. Migrate-Data enables connections to many different databases through Java Database Connectivity (JDBC) drivers available from a variety of vendors. ESF Database Convert is a toolkit that provides a user friendly step-by-step GUI. It is said that once a data source is selected and a target output is determined, then all table structure, data, schema(Oracle, SQL Server 2000 or higher, PostgreSQL), Large Text/Binary Objects (LOB), primary key/foreign key, indexes, auto-increment(serial) and default value could be easily migrated.

2. Tool Evaluation and Criteria

Quantitative evolution and Qualitative Methods are two methods used for software tools and project evaluations in software engineering [9]. Software tool evaluation also uses rating methods considering information in the program, career development process, user interaction, technical aspects of software ,and material and support services [10][11], provides a brief but a good outline of evaluation requirements for web based user interfaces.

Department of Computer Science at United States (US) Air Force Academy, evaluating three product GUI for software development related to Ada programming language [12] used a criteria containing, (i) the overall impression, (ii) ease of installation, (iii) documentation quality, (iv) ease of use and the diagnostic messages of the User Interface, and (v) ease of entering, general purpose and the adequacy of options in the GUI building. This evaluation used a ten criteria based scoring system in a scale of 1 to 10 which had 1 for very poor, 5 for "industry average" and 10 for superior. Though the hardware/platform support, coding standards, installation procedure, etc., were evaluated, no assessment had been made with respect to the performance of the product, such as efficiency and accuracy.

A software evaluation by four groups incorporating four techniques: heuristic evaluation, software guidelines, cognitive walkthroughs and usability testing [13], identified that the best is the heuristic evaluation technique. Then at second, the usability testing had performed well in finding serious problems. Guideline evaluation had been the best at finding recurring and general problems. Cognitive walkthrough technique had been roughly comparable in performance to guidelines. Although these methods are scientific and according to a standard, evaluation had been indicated as complex since it is necessary to have experienced developers to perform such evaluation

Obeidat (2006) [14], describing a survey and a statistical analysis of software evaluation criteria using a random sample of information technology professionals, indicated that the factors for software technology are efficiency, flexibility, security, language, documentation, hardware, performance, cost, reliability, availability, modularity, supplier services, and compatibility. Three most important criteria identified by information technology professionals are listed as software reliability, software performance, and software compatibility whereas costs of software, software modularity, and software language have been mentioned as the least important criteria.

Twenty three software tool evaluation criteria are introduced by some software evaluation organizations. General Description, Speech Synthesizers Supported, Installation, Memory Requirements, Navigation, Settings, Configuring, Monitoring, Alternative Cursors, Attributes, Searching, Keyboard, Macros, Pre-defined Configurations, Auto Loading, Pronunciation, Responsiveness, Applications,

Stability, Support for Sighted People, Documentation, On-line Help, Technical Support were those criteria evaluated by that group[9].

Although evaluation methods may defer on the nature of software, usage, and the functionalities, etc., the basis is the use of a marking schemes for a selected set of criteria. The design characteristics principles are considered as needs of such schema.

In case of data migration there are two important functionalities; they are 1) database connectivity which is the capability to connect, disconnect and refresh the database server, 2) database manipulation meaning the creation of tables, insertions, deletions and updating [15]. Similar to other software, in case of data migration, it is also necessary to evaluate the GUI characteristics such as its behaviour, user friendliness, accuracy achievements, and helpfulness [16].

Table 2 - Evaluation Parameters

Item	Sub Component	Main Component	Aspects
D1	Connecting	Database Connection	Database Functions Availability
D2	Refresh		
D3	Disconnecting		
D4	Creating Database	Database Manipulation	
D5	Creating Table		
D6	Inserting Data		
D7	Deleting Data		
D8	Updating Data		
D9	Querying For Data		
D10	From MySQL	Data Migration	
D11	From Others		
G1	A Separate Window for function	GUI Arrangement / Layout	GUI Functionality
G2	Availability of representative Icons		
G3	Function or a selectable menu		
G4	Use of Graphics to provide information		
G5	Availability of Known DBMS Terms	User friendliness	
G6	Absence of Confusing behavior		
G7	Capability to recover from user error		
G8	Instructions or operations follow		
G9	Absence of system errors at GUI Operations	Accuracy	
G10	Availability of indication of execution		
G11	Help support for each function	User Help	
G12	Tool tips and other supports		
G13	Additional Online Support		

3. Methodology

In the present study, the main target was to evaluate a tool for the purpose of data migration from MySQL to PostgreSQL. Based on the literature survey and consultation with the tool developers, the selection of evaluation criteria was founded on two main aspects (i) the Database functionality and (ii) the GUI characteristics. Most software enabled a user to carryout operations either with the help of the GUI or with a facility to type in specific command lines. It is known that command based functionalities require user familiarity for efficient operation while it is not required in case of a good GUI. In this work it was identified that present day tool-preference is pivoted around the GUI functionality. Cost was not considered as a factor since the objective of the evaluation was to identify a free tool. The coding aspects were also not considered as a user concern since tool modifications or customisation was not targeted. Functionality was taken as three main and several sub components while four main GUI characteristics were considered with several sub characteristics-(Table 2).

Each main component was considered equally important and hence a weighting scheme was not incorporated. Based on the efficiency, the accuracy and the user friendliness requirements each sub component was qualitatively classified using a five class scale of Very High, High, Medium, Low and Very Low. In each main category, the most important sub category was given a very high rank and then the others were given a relative ranking. In this evaluation each selected parameters were given an importance ranking but when assessing the tools, only the availability of a function and not its quality was evaluated. Therefore for each tool cells of the matrix were given either 1 or 0 value. Each qualitative ranking was given a numeric value as 5,4,3,2 and 1 for Very High, High, Medium, Low and Very Low, respectively in order to identify easily comparable evaluation indicators. For evaluation of each tool and its coverage, the cell values were converted to a numeric indicator considering the importance

with respect to both the database and GUI aspects. A simple addition of the GUI functionally rank and the Database function rank increase of available functions enabled the computation of the Component Evaluation Score (CES) at each cell and then these values were aggregated to compute the Evaluation Score of the Tool (EST). The procedure is shown by the equations 1 to 4. In this method the (CE_{ij}) Anticipated is the score that would be given to a tool if all desired functionalities are available.

This evaluation methodology described is for carrying out a preliminary assessment of tools

Database function sub component(Table 2)
$= [D_i]_{i=1-11}$ -------- (Eqn 1)
GUI functionality sub component(Table 2)
$= [G_j]_{j=1-13}$ --------- (Eqn 2)
Component Evaluation Score
$= CE_{ij} = [D_i + G_j]$ ------- (Eqn 3)
Where i=1 to 11 and j=1 to 13
Evaluation Score of Tool (EST)
$=[(\sum CE_{ij})_{Availability}/(\sum CE_{ij})_{Anticipated}]x100\%$
------- (Eqn 4)

based on the availability of capabilities and this is felt as sufficient for making a reasonable judgement. However this method requires further refinement for detailed evaluations in case there are tools that receive scores which are similar or quite close to each other. In such situations, this method requires to be expanded to cover the efficiency of operations such as time taken for data insertions, querying and carrying of the migration of a known amount of data.

4. Results and Discussion

Assessment Matrices for Database Function Availability and GUI Functionality were prepared for each of the tools selected for evaluation. As an example, Table 3 shows the assigned cell values only for the Tool "PostgreSQL plus 8.3" indicating the availability. Empty cells indicate the non availability of the user requirements. Once the values were assigned to obtain a numerical representation, they were then aggregated to obtain the EST pertaining to each main component of all evaluated tools, (Tables 4a, 4b, and 4c). Graphical representation of the Database Function Availability and EST pertaining to GUI functionality is shown in Figure 01.

The GUI Overall and Database Overall indicate the evaluation of each tool as a whole when considering the user objectives selected for this study. For the data migration purpose, PostgreSQL Plus 8.3 with a 76% EST value shows that it is much superior to the other two free tools. PostgreSQL 8.3 falls short of user expectations in the areas of GUI Arrangement and the User Help facilities especially in the area of database connection. Data Migration functionality of PostgreSQL 8.3 received a high EST value of 80% where as the other two could achieve only value of 26%.

Migrate Data 1.1 and ESF Database Convert both are having overall EST values of the same order of magnitude. The Migrate Data 1.1 shows a slightly higher value of 29% when compared with the 26% of ESF Database convert. Both these tools have shown poor GUI arrangement/layout corresponding to database manipulation and data migration as the major weakness when compared with the PostgreSQL Plus 8.3. In these two tools the said functions are executed by Execution window. Both these tools showed reasonable indicator values for the GUI facilitation of user feedback with respect to the accuracy of operation. However the database manipulation function was not very well supported by the GUI capabilities and therefore the EST values for these two tools were at a very low value of 21%. However it should be indicated that the data manipulation function in these two are carried out by specific SQL command executor. There were several other factors noted during the evaluation. In PostgreSQL plus 8.3, data migration section is designed as a separate tool section without links via a menu or buttons in the current interface. This could be considered as a strength of the tool.

In Migrate-Data 1.1, options are available only to access, modify and data migration and there was no option to create a new database or a new table within a current database. For this tool Scripts could be recommended for the handling of database tables.

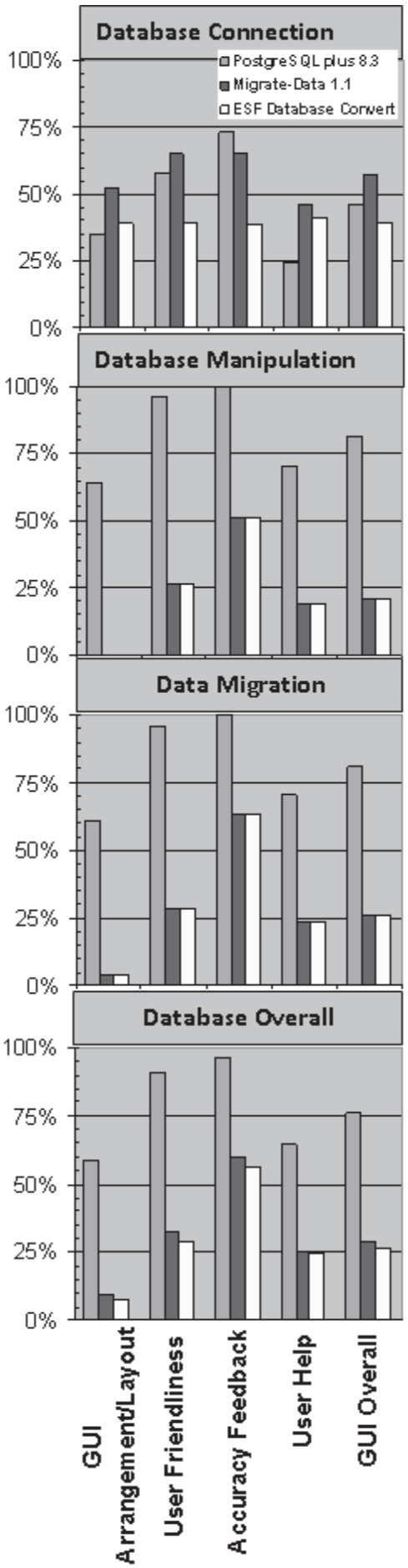

Figure 1 - EST Value Distribution

5. Conclusions

1. In case of tool selection for data migration from MySQL to PostgreSQL, it is important to consider the availability of database functions together with the GUI Functionality.
2. The present work through a critical evaluation of literature, identified 3 main and 11 sub database functions together with 4 main and 13 sub components of GUI functionality, as important parameters for MySQL to PostgreSQL data migration tool evaluation
3. A qualitative ranking of parameters with the use of a set of representative numerical indicators, together with a simple aggregation technique enabled to evaluate the strengths and the weaknesses of the free tools available for data migration from MySQL to PostgreSQL
4. The study showed a simple and a systematic methodology to carryout a data migration tool evaluation which is a very important requirement in the context of the present day where one finds a significant number of tools available over the internet
5. For the Data migration purpose, PostgreSQL plus 8.3 with a 76% EST value indicated its superiority over the other two namely, the Migrate Data 1.1 and ESF Database Convert, which received only 29% and 26% as EST values respectively

6. References

1. MySQL, http://en.wikipedia.org/wiki/MySQL visited 10th December 2010
2. PostgreSQL http://en.wikipedia.org/wiki/PostgreSQL Visited 17th December 2010.
3. Unix, http://unix.org.in/2010/10/postgresql-vs-mysql/ Visited 17th December 2010.
4. Web Masters, http:// http://www.webmasters.com/doit.htm Visited 17th December 2010.
5. quality web Programming, http://www.quality-web-programming.com/blog/use-of-mysql-and-php-for-web-programming/ Visited 17th December 2010.

6. Ivens R.J.C., ROCLASI Software Solutions, The Netherlands, June 26 2008 download.roclasi.com/PgSQL_vs_MySQL.pdf Visited 10th December 2010.

7. Hunter S.R MySQL vs. PostgreSQL, http://articles.techrepublic.com.com/5100-10878_11-1050671.html 01st July 2001.

8. "How to Make Proper Migration" , http://wiki.postgresql.org/wiki/How_to_make_a_proper_migration_from_MySQL_to_PostgreSQL visited 05th January 2011.

9. David Andrews , GUI Evaluation Criteria, 13th July 1994 http://www.nfbcal.org/nfb-rd/0300.html visited 5th January 2011.

10. Software Evaluation Criteria , http://associationdatabase.com/aws/NCDA/asset_manager/get_file/3404 visited 09th January 2010.

11. Javanovic J. Designing User interface for Business Web Applications www.smashingmagazine.com/2010/.../designing-user-interfaces-for-business-web-applications/ visited 09th January 2010.

12. Robert Steigerwald, David cook and Leslie Dupaix , An Evaluation of Graphical user interface Builders for Ada Software Development, 2003.

13. Robin Jeffries, James R. Miller, Cathlen Wharton and Kathy M.Uyeda (1991), User Interface evaluation in the real world: A comparison of four techniques, Hewlett Packard Laboratories. www.hpl.hp.com/techreports/91/HPL-91-03.pdf Visited 19th December 2010]

14. Obeidat, M. A , Evaluation of New Software Technology: an Empirical Study, International Management Review, Vol. 2 No. 3 2006.

15. Query, Edit, Browse and Manage databases http://www.razorsql.com/features/sql_gui.html visited 10th December 2011.

16. Ian sommerville , Software Engineering, 6th Edition 2002.

Table 3 - Assessment Matrix for Database Function Availability and GUI Functionality of PostgreSSQL plus 8.3

	Data Base Function Availability >		Database Connection			Database Manipulation						Data Migration	
			D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11
GUI Functionality >		Rank	VH	H	L	H	VH	VH	VH	VH	VH	VH	L
GUI Behaviour	G1	VH	1			1	1	1			1		1
	G2	M	1	1			1		1	1	1		1
	G3	VH				1	1		1	1	1		1
	G4	M		1			1				1		1
User Friendliness	G5	VH	1	1		1	1	1	1	1	1	1	1
	G6	H	1	1		1	1	1	1	1	1	1	1
	G7	VH	1	1		1	1	1	1	1	1	1	1
	G8	M				1	1	1		1	1	1	1
Accuracy Feedback	G9	VH	1	1		1	1	1	1	1	1	1	1
	G10	VH	1	1		1	1	1	1	1	1	1	1
User Help	G11	VH				1	1	1	1	1	1	1	1
	G12	M	1	1		1	1	1	1	1	1	1	1
	G13	M											

Table 4 a - EST Values for Tool and for Main Components: PostgreSQL plus 8.3

	Database Connection	Database Manipulation	Data Migration	Database Overall
GUI Arrangement/ layout	35%	64%	61%	58%
User Friendliness	58%	96%	96%	91%
Accuracy Feedback	73%	100%	100%	96%
User Help	25%	70%	71%	64%
GUI Overall	46%	81%	80%	76%

Table 4 b - EST Values for Tool and for Main Components: Migrate Data 1.1

	Database Connection	Database Manipulation	Data Migration	Database Overall
GUI Arrangement/layout	52%	0%	4%	10%
User Friendliness	65%	26%	28%	33%
Accuracy Feedback	65%	51%	63%	60%
User Help	46%	19%	24%	25%
GUI Overall	57%	21%	26%	29%

Table 4 c - EST Values for Tool and for Main Components: ESF Database Converter

	Database Connection	Database Manipulation	Data Migration	Database Overall
GUI Arrangement/layout	39%	0%	4%	8%
User Friendliness	39%	26%	28%	29%
Accuracy Feedback	38%	51%	63%	56%
User Help	41%	19%	24%	24%
GUI Overall	39%	21%	26%	26%