

Database and Flat File Validation Tool

Fathima Gihan K.S.¹, Aneesh M Haneef²

¹M.Tech Student, M.E.S College of Engineering, India

²Assistant Professor, M.E.S College of Engineering, India

ABSTRACT : *New business models, constant technological progress, as well as ever-changing legal regulations require that companies replace their business applications from time to time. As a side effect, this demands for migrating the data from the existing source application to a target application. Since the success of the application replacement as a form of IT maintenance is contingent on the underlying data migration project, it is crucial to accomplish the migration in time and on budget. Migration of database-to-database and flat file to database is complicated. Major challenge on quality assurance is to make sure no data has been lost or altered. To achieve this business goal an integrated many to many database and file comparison tool is required. In this research work we propose a multi-database and flat file validation tool to make sure quality assurance on migration projects.*

Keywords - Database comparison, Data migration, Flat file validation, Quality assurance, Testing.

I. INTRODUCTION

1.1. Data Migration Process

Nowadays, database technology has become the core technology of information system in each organization. Along with development of information technology, database technology has also been promoted, which certainly includes database management system technology. Since the appearance of these advancements, old systems have already been replaced by more powerful systems. During the process of updating, data migration is a significant issue. The main objective that all migration projects have in common is to permanently move the data from the source business application into the target in implementing a specific process.

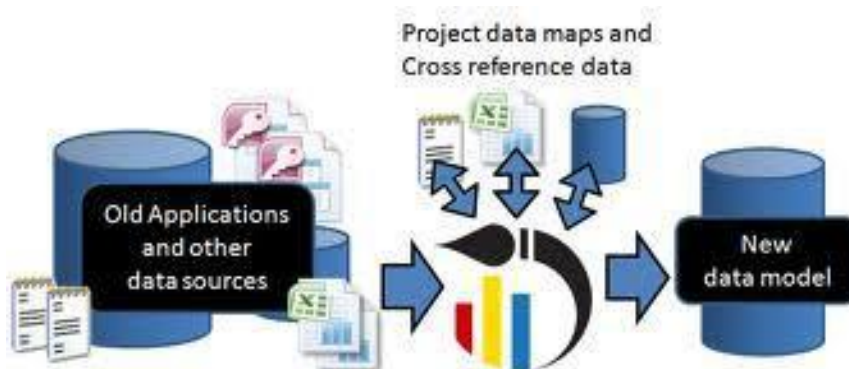


Fig.1: Data Migration Process

1.2. Problem Statement

Existing backend systems may have multiple database and files. Each database holds enough number of tables and massive data records in each table. As part of the migration projects, source database table structure would be splitting across to different tables in destination database based on their new design. The major challenge on quality assurance is to make sure no data has been lost or altered. Data validation tools available in the market are limited to one to one database comparison. To achieve the business goal an integrated many to many database and file comparison tool is required.

II. PROBLEM ANALYSIS

2.1. Database to Database Validation

The main objective that all migration projects have in common is to permanently move the data from source business application to the target application by implementing a specific process. Here data validation on multiple databases will be the major challenge and is very difficult to achieve. We have number of database comparison tools available in the industry, but they are limited to one-to-one comparison. Using these tools we can compare multiple tables within same type of database as mentioned in the below Fig. 2 [1][2].

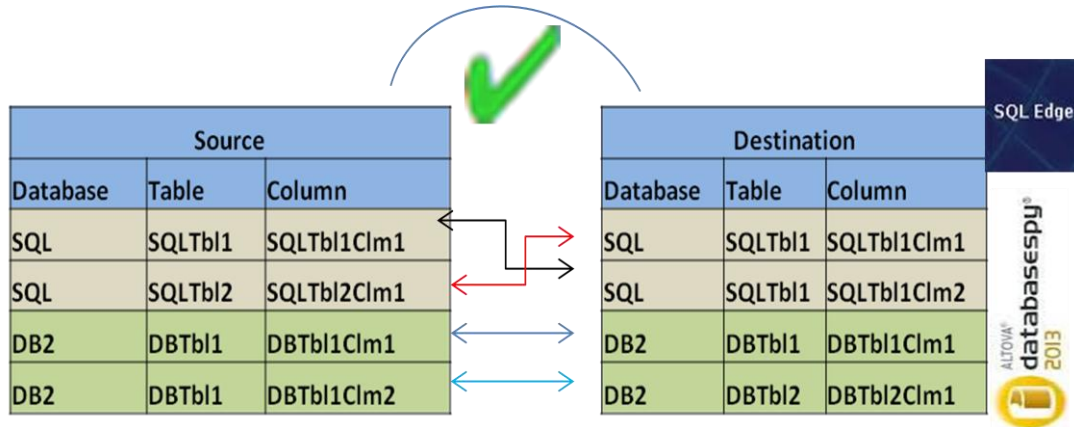


Fig.2: Similar type of database validation

The first problem defines in the below requirement for data validation on multiple tables within multi-databases as shown in Fig. 3. Here the source and destination databases are in different types and the source table data may be scattered on different databases tables in destination. The key challenge is to make sure no data has been altered or lost during the migration.

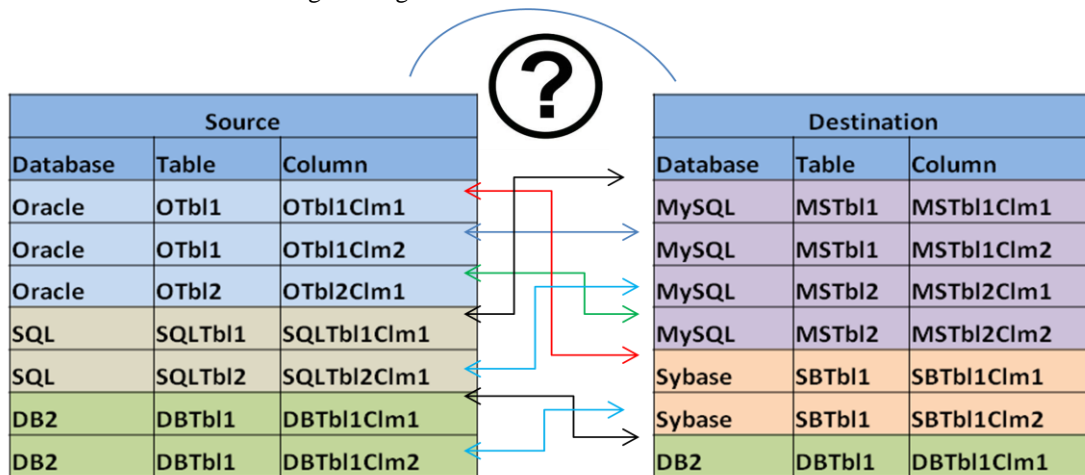


Fig.3: Multi type database validation

2.2. Flat Files to Flat File Validation

Flat files are text files in which record columns are separated based on individual file formats. As shown in the below Fig.4, Flat file A and B are having similar file format and similar record order. These flat file contains five columns such as name, phone number, sex, address and age. The data values are placed in the files based on the each column position in the format of the file. Data validation of the below requirement will be achieved though line by line comparison on source and destination data files. There are number of tools available in the industry which can do these operations.

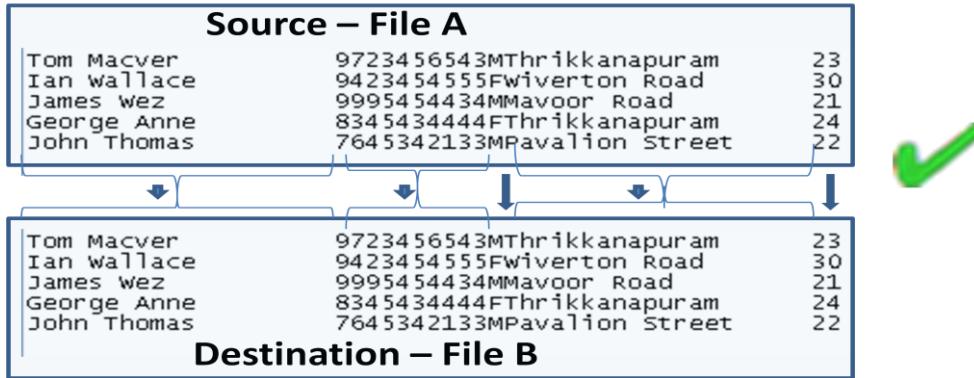


Fig.4: Flat File validation with similar file format and record order

As shown in Fig. 5, the requirement is to validate Flat File A and B where file A and B are having similar file format but different record order. Line by line comparison of file will be applicable here as well with a record order constraint. For validation of this requirement ‘Beyond Compare’ tool would be the best choice.

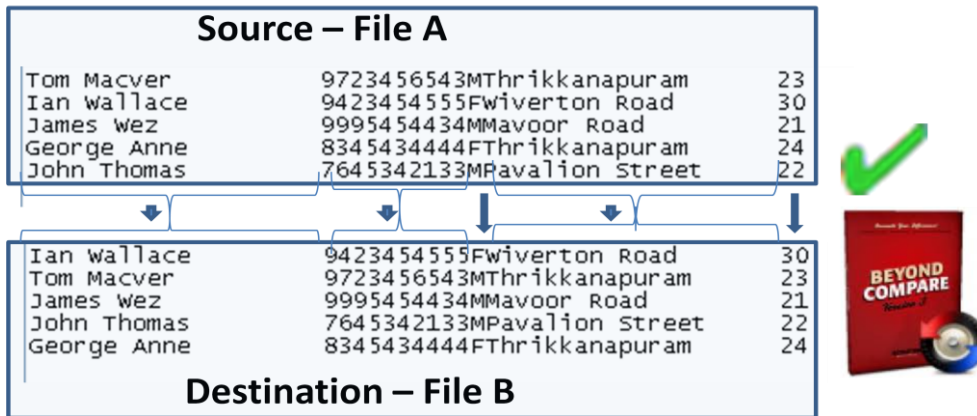


Fig. 5: Flat File validation with similar file format and different record order

The second problem defines as shown in Fig. 6. The requirement is to validate flat file A and B when the source and destination flat files which is having different file formats and may have different record order as well.

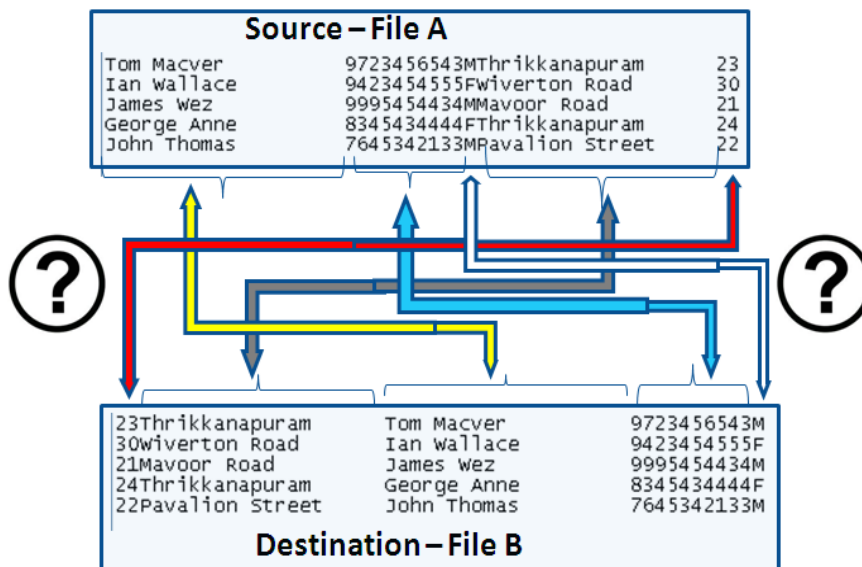


Fig. 6: Flat File validation with different file format

2.3. Database to Flat File Validation

The third problem define from the below (Fig. 7) requirement to validate the data values on source Flat file and destination databases. Source file data would be shifted to multiple databases and the data might be placed in multiple tables. Key challenge here is to validate the flat file columns with different table values on multiple databases.

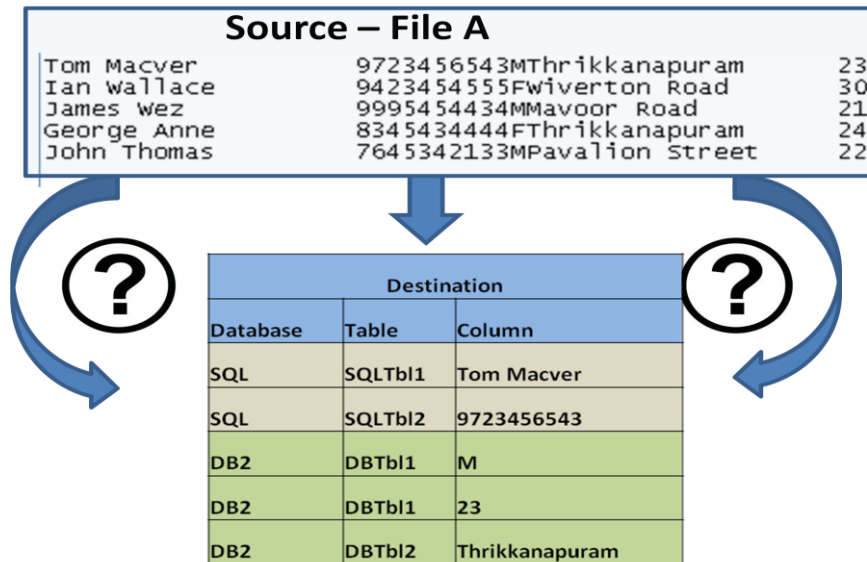


Fig. 7: Validation of Flat File with Databases

2.4. Analysis on major database tools

Conducted a survey on major database tools and the analysis results are detailed below. We can find many tools for validation of data on multiple tables within a same database, but data validation between multiple databases in the key challenge [3][4][5][6].

Tools	License	Windows	Linux	Query Generation	Data Extraction	DB Compare - One to One	DB Compare - One to Many	DB Compare - Many to Many	DB Compare - with Flat File	File Comparison
ABIS-Query	Proprietary	✓	✗	✓	✓	✗	✗	✗	✗	✗
SQL Compare 10	Proprietary	✓	✗	✗	✗	✓	✗	✗	✗	✗
SQL Delta	Proprietary	✓	✗	✗	✗	✓	✗	✗	✗	✗
ACDB	Shareware	✓	✓	✓	✓	✗	✗	✗	✗	✗
Adminer	Apache license	✓	✓	✓	✓	✗	✗	✗	✗	✗
Advanced Query Tool	Proprietary	✓	✗	✓	✓	✗	✗	✗	✗	✗
Aqua Data Studio	Proprietary	✓	✓	✓	✓	✓	✗	✗	✗	✓
DatabaseSpy	Proprietary	✓	✗	✓	✓	✓	✗	✗	✗	✗
DbSchema	Proprietary	✓	✓	✓	✓	✗	✗	✗	✗	✗
SQL Edge	Proprietary	✓	✓	✓	✓	✗	✗	✗	✗	✗

Fig. 8: Analysis on major database tools

III. PROPOSED SOLUTION

Software testers often come across instances in which data from one database is moved to another, data from different databases are joined to put in a single table in another database and they need to validate the data. The validation of the migration is usually a really tedious and time crunching task. Another side is manual operations using flat files is a nightmare for every data stage and mainframe testers. Here is where an automation tool comes into play [7]

We are proposing a new automation tool ‘Database and Flat File Validation Tool’ for Multi-database and Flat File validation. This tool would be executed on multiple databases and files together - source and destination database can be different from each other (E.g. source is DB2 & SQL and destination is Oracle &

MySQL). The validation tool would be executed although the source table fields are scattered on different tables in destination database. The tool would include more flexible GUI to accept end user inputs and it would have an excellent reporting mechanism. The tool would be capable to generate the required complex SQL queries dynamically based on the end-user input on frontend screen.

Quality assurance is a major challenge on data migration process. All data records on source and destination databases and flat files should be validated correctly to make sure no data has been lost or altered. Manual validation process cannot accommodate within the project time line and budget. Database and Flat File Validation Tool makes this simpler and accurate. [8]

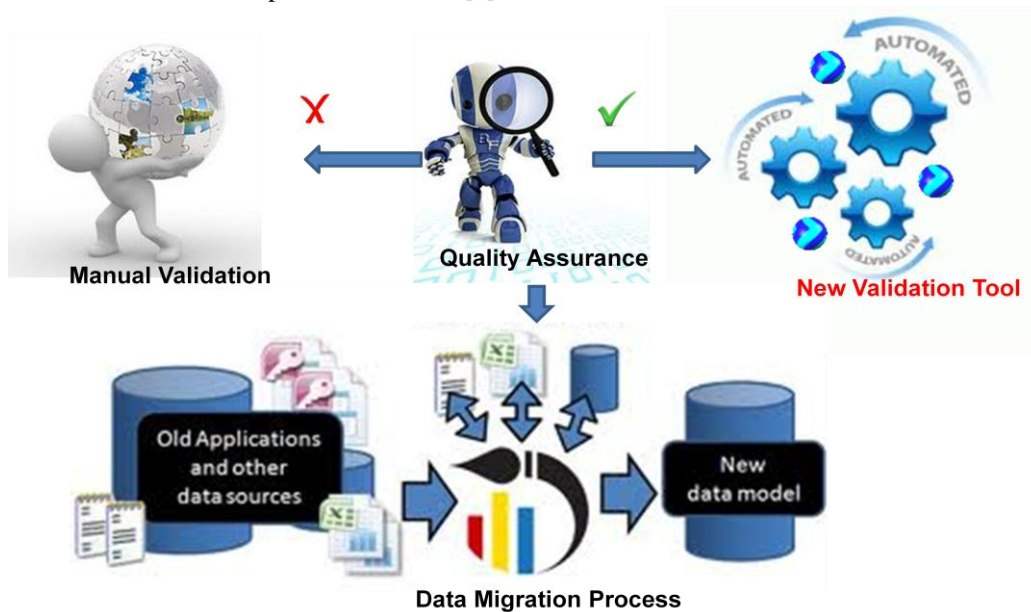


Fig. 8: Scope of new validation tool

IV. CONCLUSION

Major challenge on data migration quality assurance is to make sure no data has been lost or altered. Data validation tools available in the market are limited to one to one database comparison. We proposed a new automation tool ‘Database and Flat File Validation Tool’ for Multi-database and Flat File validation. The tool would be executed on multiple databases and files together - source and destination database can be different from each other. The validation tool would be executed although the source table fields are scattered on different tables in destination database. The tool would include more flexible GUI to accept end user inputs and it would have an excellent reporting mechanism. The tool would be capable to generate the required complex SQL queries dynamically based on the end-user input on frontend screen.

REFERENCES

- [1]. Jun Li Hewlett Packard Labs., Palo Alto, CA, USA Stephenson, B. ; Motahari-Nezhad, H.R. ; Singhal, S., “*GEODAC: A Data Assurance Policy Specification and Enforcement Framework for Outsourced Services*” Pages: 340–354, Journals & Magazines, Services Computing, IEEE Transactions on Oct. – Dec. 2011.
- [2]. Elamparithi, Pollachi, India, ‘*Database Migration Tool (DMT) – Accomplishments & Future Directions*’, Communication and Computational Intelligence (INCOCCI), 2010 International Conference on 27-29 Dec. 2010.
- [3]. <http://www.abis-query.com/>
- [4]. <http://www.aquafold.com/>
- [5]. <http://www.red-gate.com/products/sql-development/sql-compare/>
- [6]. <http://www.sqldelta.com/>
- [7]. Guoliang Li Dept. of Comput. Sci. & Technol., Tsinghua Univ., Beijing, China Jianhua Feng ; Chen Li, “*Supporting Search-As-You-Type Using SQL in Databases*” Pages: 461–475, Journals & Magazines, Services Computing, IEEE Transactions on Feb. 2013.
- [8]. Matthes, F. Software Eng. for Bus. Inf. Syst., Tech. Univ. Munchen, Munich, Germany Schulz, C. ; Haller, K. “*Testing & quality assurance in data migration projects*” IEEE International Conference on 25-30 Sept. 2011.