



SYNASC 2010

Published by:



WoSS

1st Workshop on Software Services: Frameworks and Platforms

e-System for Automatic Data Migration

Andreea Marin, Ciprian Dobre, Decebal Popescu, Valentin Cristea

Emails: andreea.marin@cti.pub.ro, {ciprian.dobre, decebal.popescu, valentin.cristea}@cs.pub.ro

University “*Politehnica*” of Bucharest
Faculty of Automatic Control and Computer Science

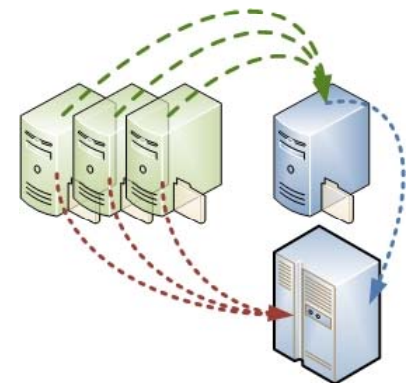
Outline

- Scope and motivation
- Data Migration
- e-System for Automatic Data Migration
- Application Architecture
- Implementation Details
 - User Interface
 - Rule Mapping Module
 - Data Conversion Module
- Case study
- Conclusion and future work



Scope and Motivation

- Companies tend to develop in size - old storage facilities evolve into complex data storage systems
- Many e-Systems are made by composing services that use data stored differently
- Data migration - translation of data from one format to another format or from one storage device to another storage device
- Data migration is expensive
 - \$5 billion on data migration, considering software services and consulting
- Organizations tend to collect serious amounts of data



Data migration approaches

- Talend Open Studio
 - data integration and business modeling software product
 - + transfer data between any two sources
 - + support for the most important database systems and for the best known file types.
 - schemas of the source and destination entities have to be known at design time
- SQL Server Management Studio Express,
 - + workflows for data transfer
 - tasks that can be executed on top of the database.
 - transfers data in known file types (csv, xls) or in Oracle and SQL Server Databases having *the same* schema.
 - the transfer depends on the Oracle Database System's version (10g, 9i).
- Oracle SQL Developer
 - transfers data in flat files

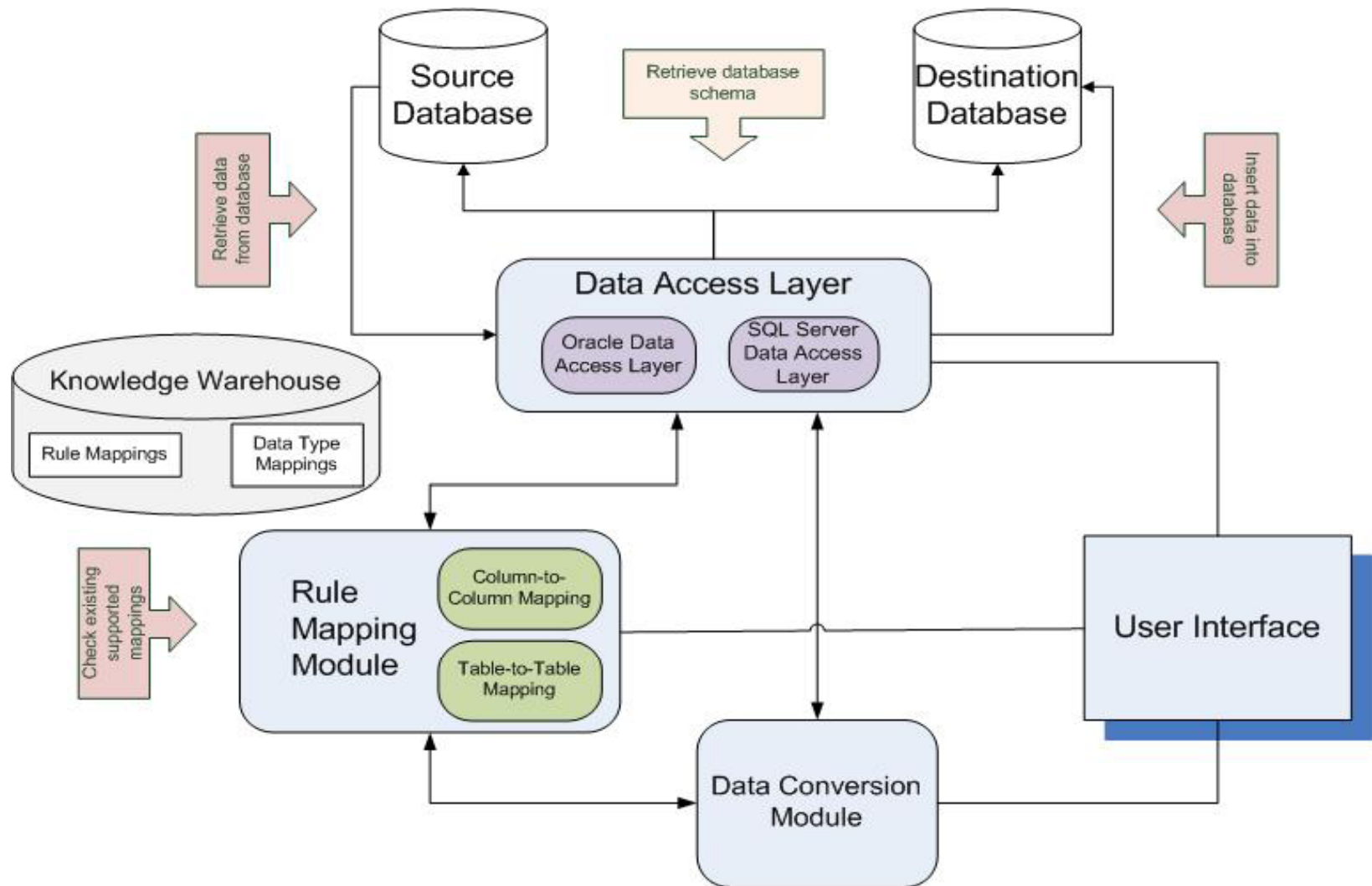




e-System for Automatic Data Migration

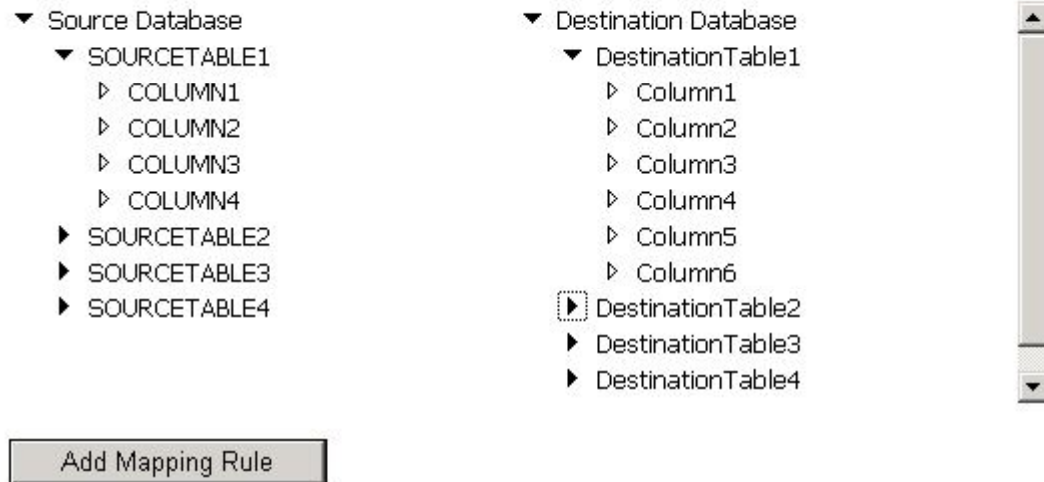
- Data mapping between two sources
- Dynamic data transfer
 - No prior details about the structure of the data
- The user has full control over the data migration process
- Mechanisms to
 - Automatic recognize data types
 - Map data fields belonging to different data sources
 - Move data on triggered events
- Specialized modules for
 - Rule mapping
 - Data conversion
 - Data access

Architecture



User Interface

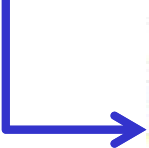
- Authentication
- Creation and management of mapping rules
- Display of data schemas, conflicts or resolutions
- Display of migration results



Rule Mapping Module

- Repository of mapping rules
- Allows management of created mapping rules
 - Rule creation and registration
 - Rule modification
 - Rule deletion
- Checks the validity of the rules

Mapping rules



	SourceTable	SourceColumn	DestinationTable	DestinationColumn
<u>Delete</u>	SOURCETABLE1	COLUMN1	DestinationTable1	Column2
<u>Delete</u>	SOURCETABLE2	COLUMN3	DestinationTable3	Column1
<u>Delete</u>	SOURCETABLE4	COLUMN9	DestinationTable1	Column6



Rule Mapping Module

- Column-to-column mappings
 - The user chooses a column from a table from the source database and a column from a table in the destination database to create a mapping
- Table-to-table mappings
 - Tables from the source and destination databases are chosen to create a mapping
- Knowledge Warehouse
 - Predefined rules that can help the mapping process
- Management of user generated errors
 - E.g. the user might create a table-to-column mapping



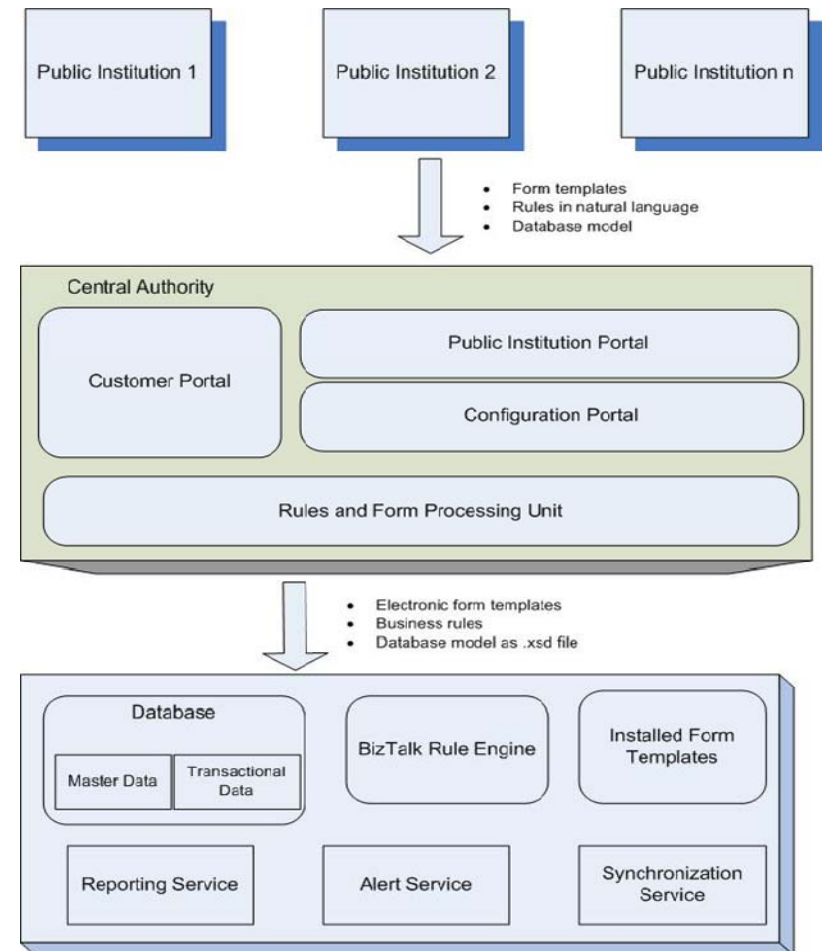
Data Conversion Module

- Assists the rule mapping and migration functions with data conversions
- Implicit data conversions
 - Conversions between well-known database data types
 - E.g. conversions between SQL Server and Oracle such as Binary to Raw, Image to Long
 - Predefined rules
- Explicit data conversions
 - Conversions between non-compatible data types
 - E.g. numerical data types into strings
- The user is assisted by the system throughout the migration process

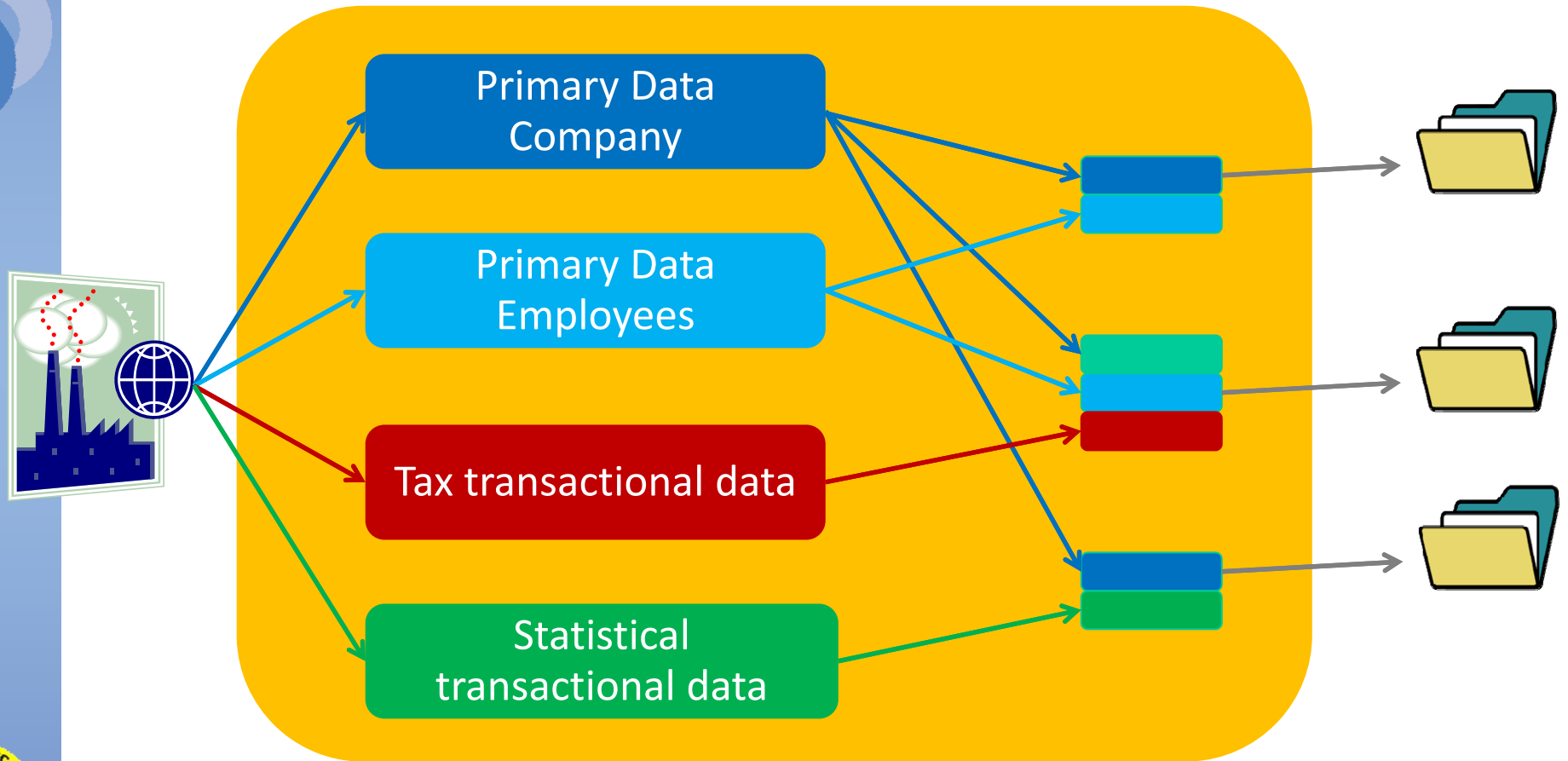


Case Study (I)

- e-Services system for public administration reporting services
- Central Authority
 - Management of rules, reporting templates, processes
- Processors
 - Automatic generation of reports



Aggregation and sending statements



Case Study (II)

- The e-Services system:
 - Improving management of routine tasks
 - Automatically managing interactions between the organization and the public administration
- The data migration software provides the initial data for the reporting system.
 - The data resides in the user's own database
- Pilot implementation focuses on **010 Fiscal Registration Declaration** = declaration of amendments for judicial persons, associations, and other entities without judicial personality
- Specific Romanian public administration reports
 - Social insurance
 - Environmental
 - Fiscal reporting
 - Statistical data



Conclusions

- System designed to automate the data migration between various data sources
 - No prior knowledge about the data schema
 - No data schema has to be defined at design time – increased usability for non-technical users
- The application supports migration between similar and different databases.
 - Data migration between two SQL Server databases.
 - Migration between an Oracle database instance and a SQL Server.
- As a case study, we evaluated the solution in a real-world system implementation designed to automatically collect report data.
 - The migration system optimized the performance by ensuring a transparent data retrieval process.
 - It allows the user to easily manage the data mapping processes for filling the required reports.



Q&A

Thank you! 😊

