Toward Propagating Data Warehouse Evolution on Data Marts

Saïd Taktak and Jamel Feki

University of Sfax,
FSEGS Faculty, P.O.Box 1088
Miracl Laboratory, Sfax, Tunisia
{said.taktak,jamel.feki}@fsegs.rnu.tn

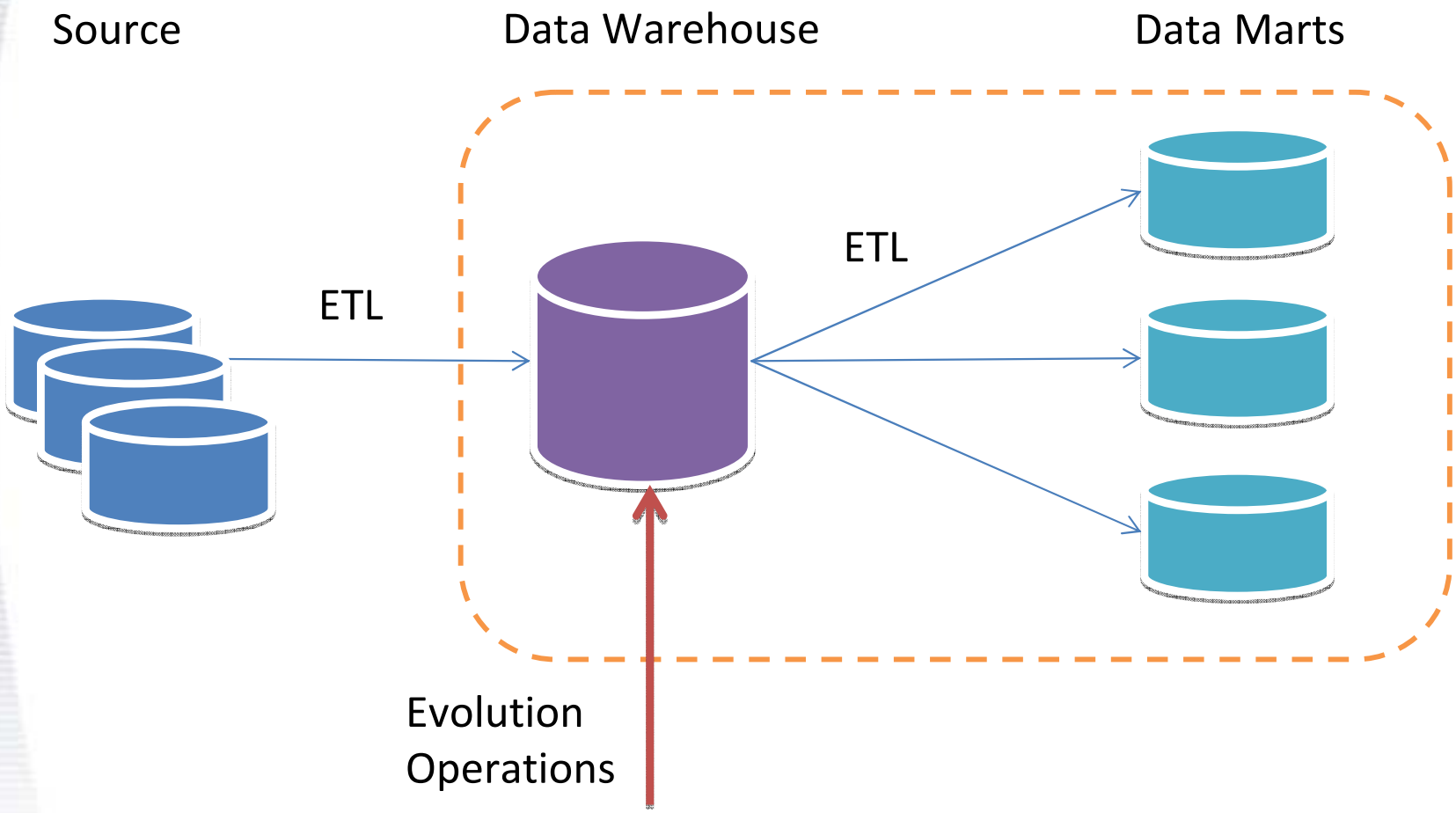
medi 2012
October 3rd & 5th 2012
Poitiers, Futuroscope - FRANCE
INTERNATIONAL CONFERENCE
MODEL & DATA ENGINEERING

Outline

- Context and Motivation
- Related works
- Approach for Propagating DW Evolution
- Transformation Rules
- Conclusion

Context and Motivation

- *Contexte and motivation*
- Related Works
- Approach for Propagating DW Evolution
- Transformation Rules
- Conclusion



Context and Motivation

- The DW is dynamic :
  - Data are periodically loaded
  - The conceptual schema may evolve
- Difficulty to definitively determine the DW schema during the design phase (Kimball) :
  - Incompleteness of needs initially captured during the design phase
  - Evolution of decision makers needs
  - Evolution/New business processes

Related Works

- *Schema Versioning*
  - aims to keep track of changes made to different versions of the *DW* schema
- *View Maintenance*
  - assumes that a DW is a set of materialized views built/loaded directly from the IS data source which require to be maintained
- *Schema Evolution*
  - adopts the hypothesis of unique schema. It considers that the schema can have only one version at a time: the current version in which all the DSS data are stored

Related Works

Comparison of DW evolution approaches

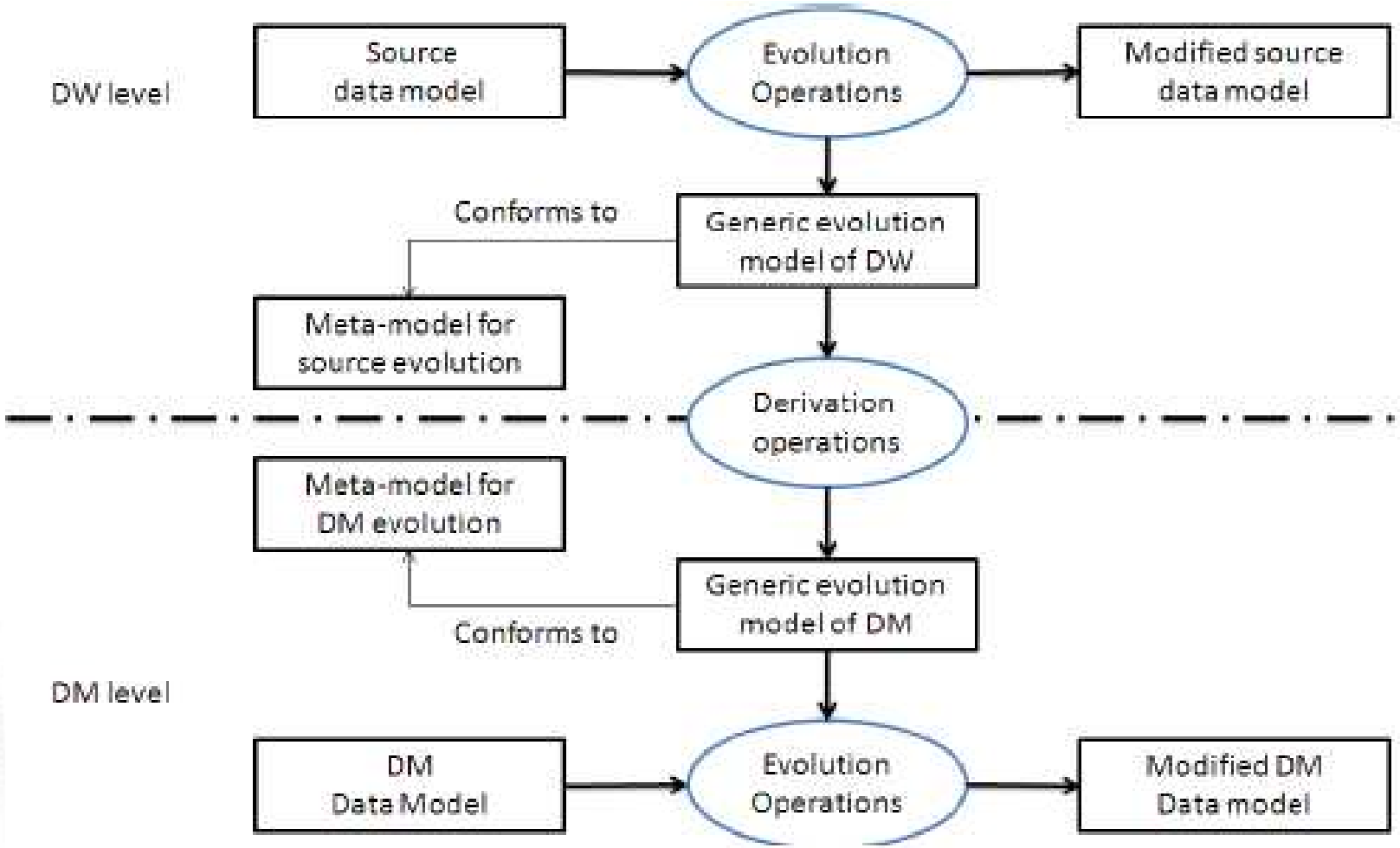
		Hurtado et al. (1999)	Blaschka et al. (1999)	Benitez-Guerrero et al. (2004)	Favre et al. (2007)	Papastefanatos et al. (2009)
Evolution of DW schema	Dimension	✓	✓	✓	✓	
	Hierarchy	✓	✓	✓	✓	
	Fact table		✓	✓		
Instances evolution		✓				
Materialized views evolution		✓				
Cubes evolution				✓		
ETL evolution						✓
Conform to a Meta-model			✓		✓	
Software prototype			FIESTA	WHES	WEDriK	HECATAEUS

Related Works

- Recapitulation:
  - All evolution strategies are at a single modeling level
    - Schemas before and after changes are expressed conforming to the same meta-model
  - Few investigations in DW/DMs schema evolution
    - DW is a conventional database.
    - DMs are multidimensional modeled

Approach for Propagating DW Evolution

- Objective: study the impact of the DW schema changes on its DMs.



Approach for Propagating DW Evolution

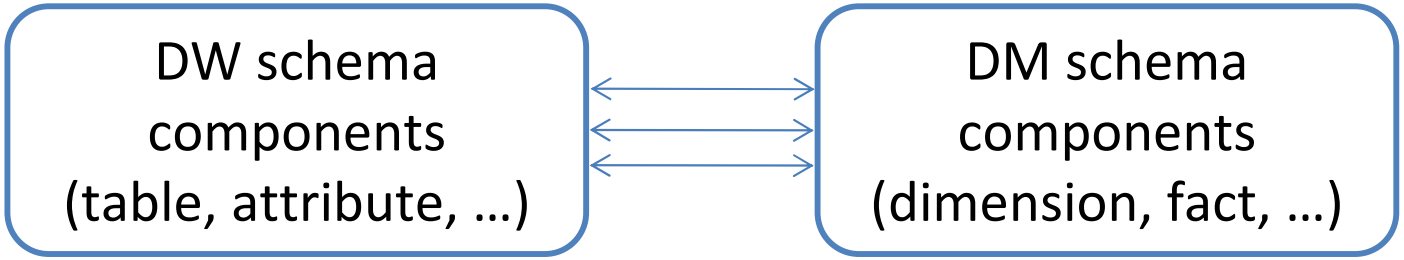
- DW Evolution Operations :
  - affect the DW schema.
  - two categories::
    - Basic operations
      - addition, deletion and modification of attributes, tables and constraints
    - Composite operations
      - composition of basic operations (splitting table, …)

Approach for Propagating DW Evolution

- DM Evolution Operations :
  - affect DM schema in order to achieve the propagation of DW changes
    - Add dimension
    - Append level
    - Insert level
    - Add fact
    - …

Approach for Propagating DW Evolution

- DW-DM Mapping :
  - establishes the correspondances between :



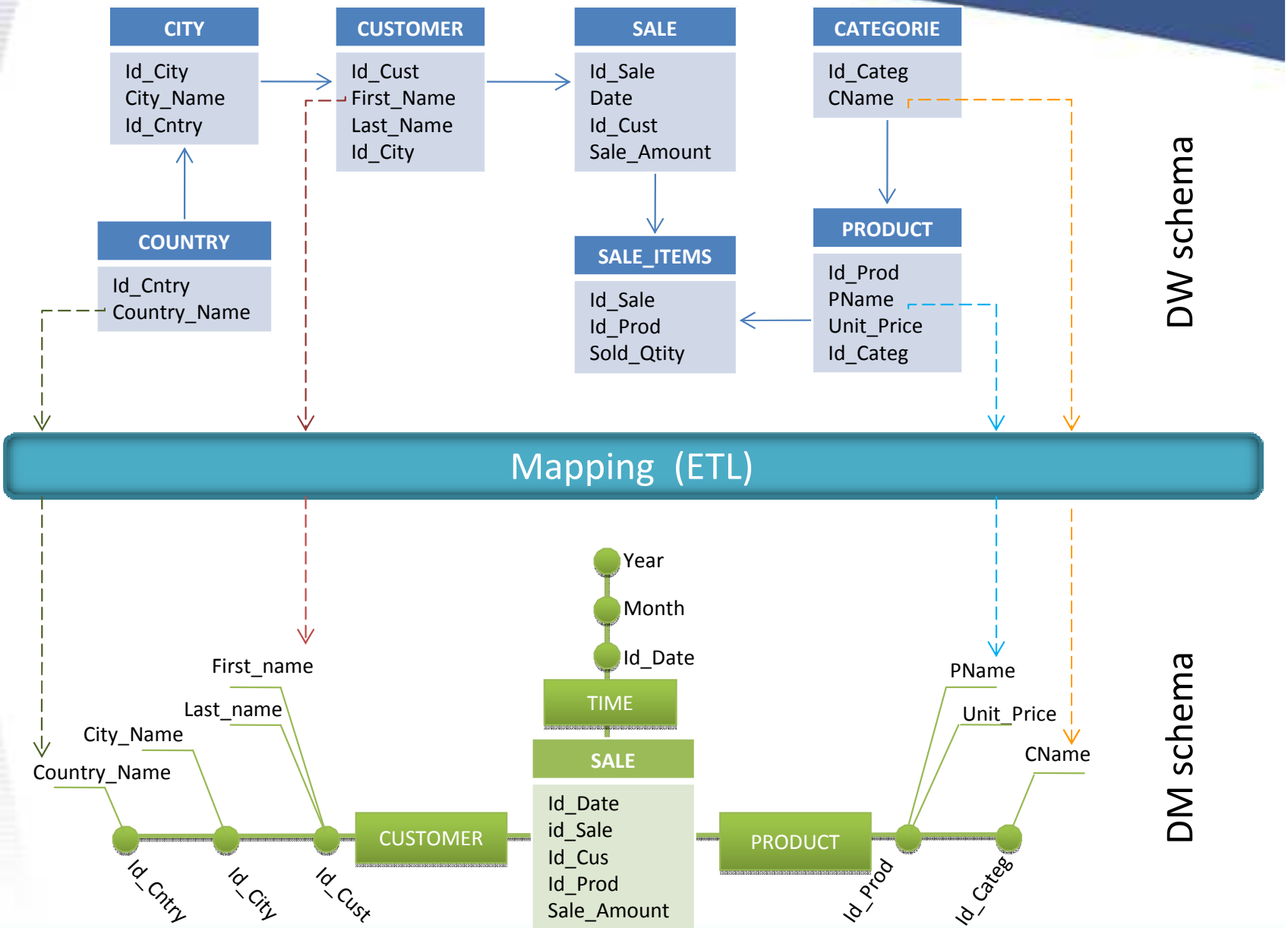
  - define mapping functions
  - these functions require an access to the mapping metamodel

Transformation Rules

- Evolution Rules
  - A table *T* added to the DW can play different roles :
    - *T creates a new dimension*
    - *T appends a terminal level to an existing hierarchy*
    - *T inserts a new hierarchy level*
    - *T creates a new fact table*

Transformation Rules

- Contexte and motivation
- Related Works
- Approach for Propagating DW Evolution
- ***Transformation Rules***
- Conclusion

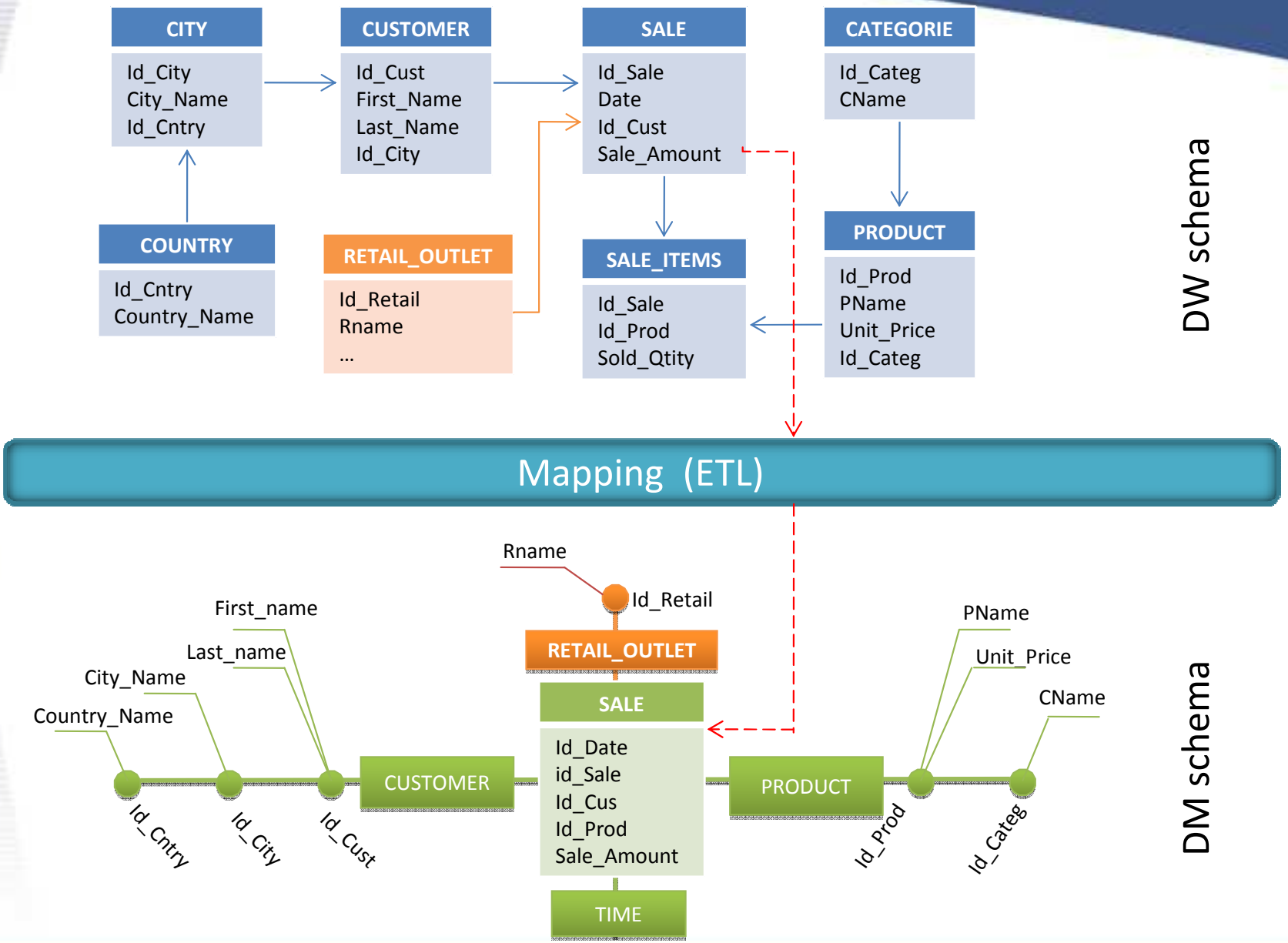


Transformation Rules

- Case (1) : *T creates a new dimension:*
  - *Condition*
    - If *T added to the DW is referenced by one table that* feeds a fact *F*
  - *Result*
    - *T feeds the new dimension for F.*
    - *The attributes of D_T are issued from T.*

Transformation Rules

- Contexte and motivation
- Related Works
- Approach for Propagating DW Evolution
- ***Transformation Rules***
- Conclusion

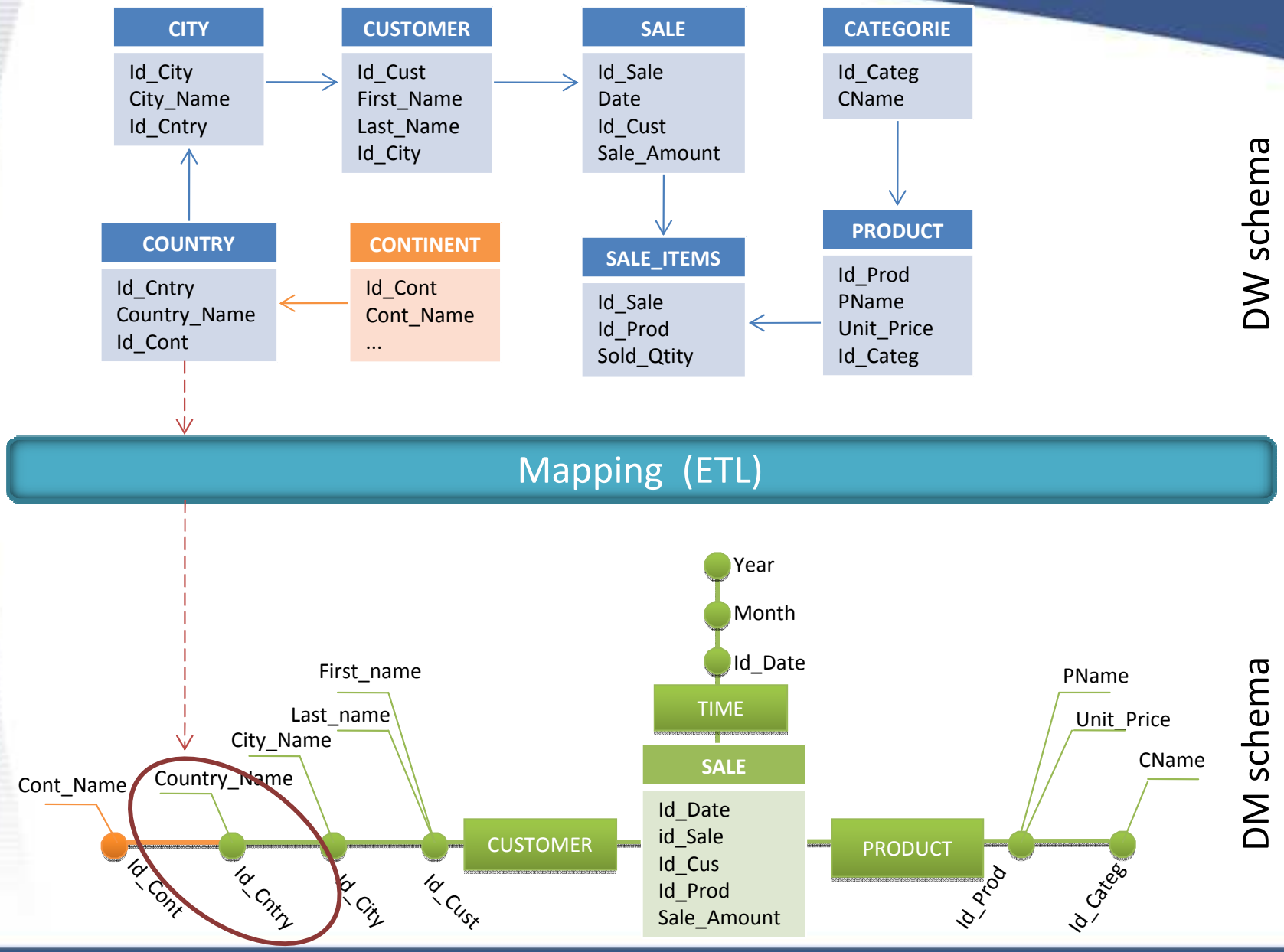


Transformation Rules

- Case (2) : *T appends a terminal level to an existing hierarchy*
  - *Condition*
    - If the table *T added to the DW is referenced by* n *tables those* feed dimensions
    - *T does not refer to any table*
  - *Result*
    - *T can feed a new level* of hierarchy in each of these dimensions
    - *The identifier of T will be a terminal parameter in multiple hierarchies*
    - *Textual attributes of T become weak attributes for the added level*

Transformation Rules

- Contexte and motivation
- Related Works
- Approach for Propagating DW Evolution
- *Transformation Rules*
- Conclusion

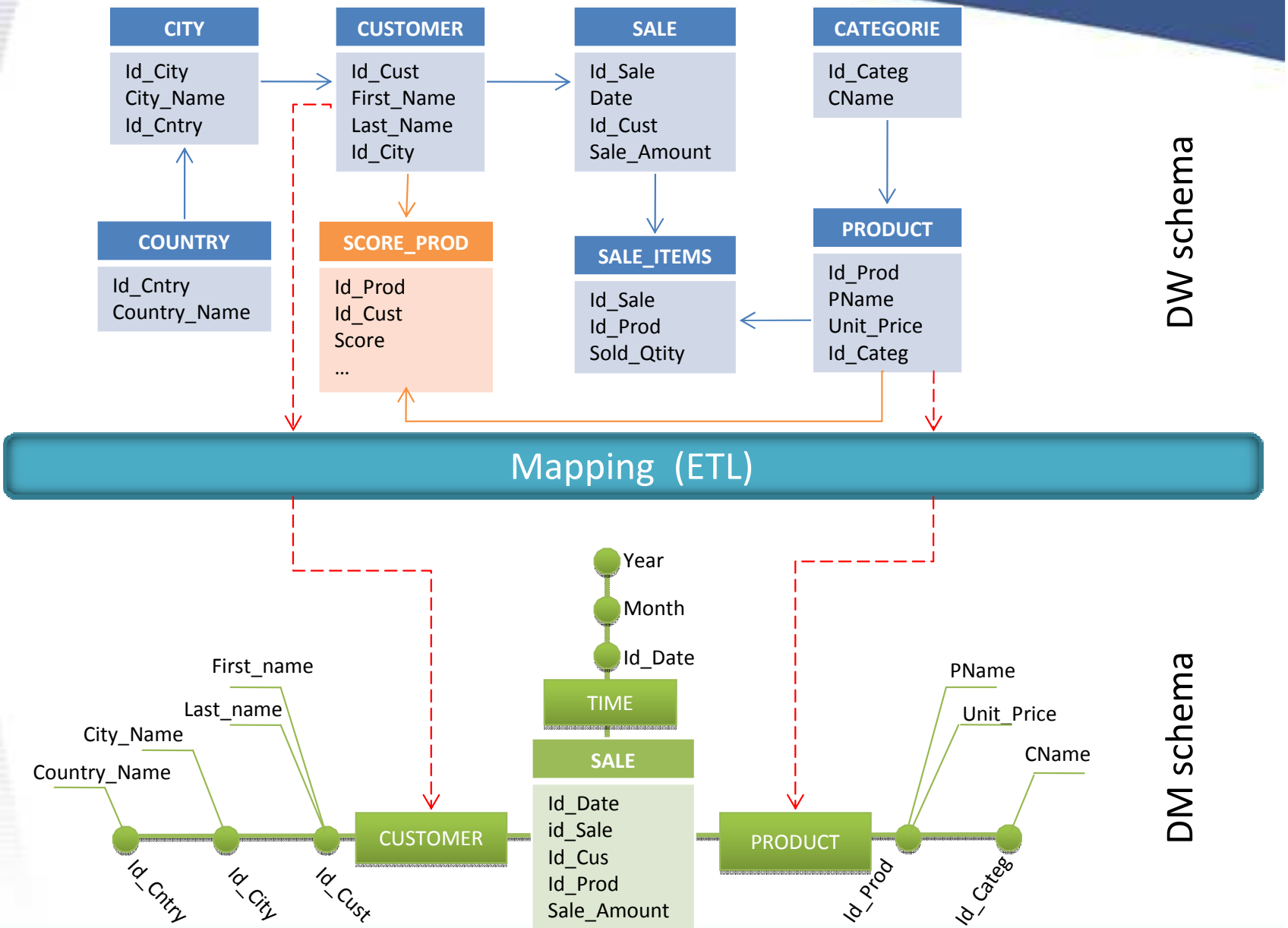


Transformation Rules

- Case (3) : *T creates a new fact table*
  - *Condition*
    - If table T is not referenced by any table of the DW
    - T references several tables loading different dimensions of the DM
    - *T has numeric attributes*
  - *Result*
    - *T is a plausible fact*
    - *numeric attributes of T become measure for the added fact*

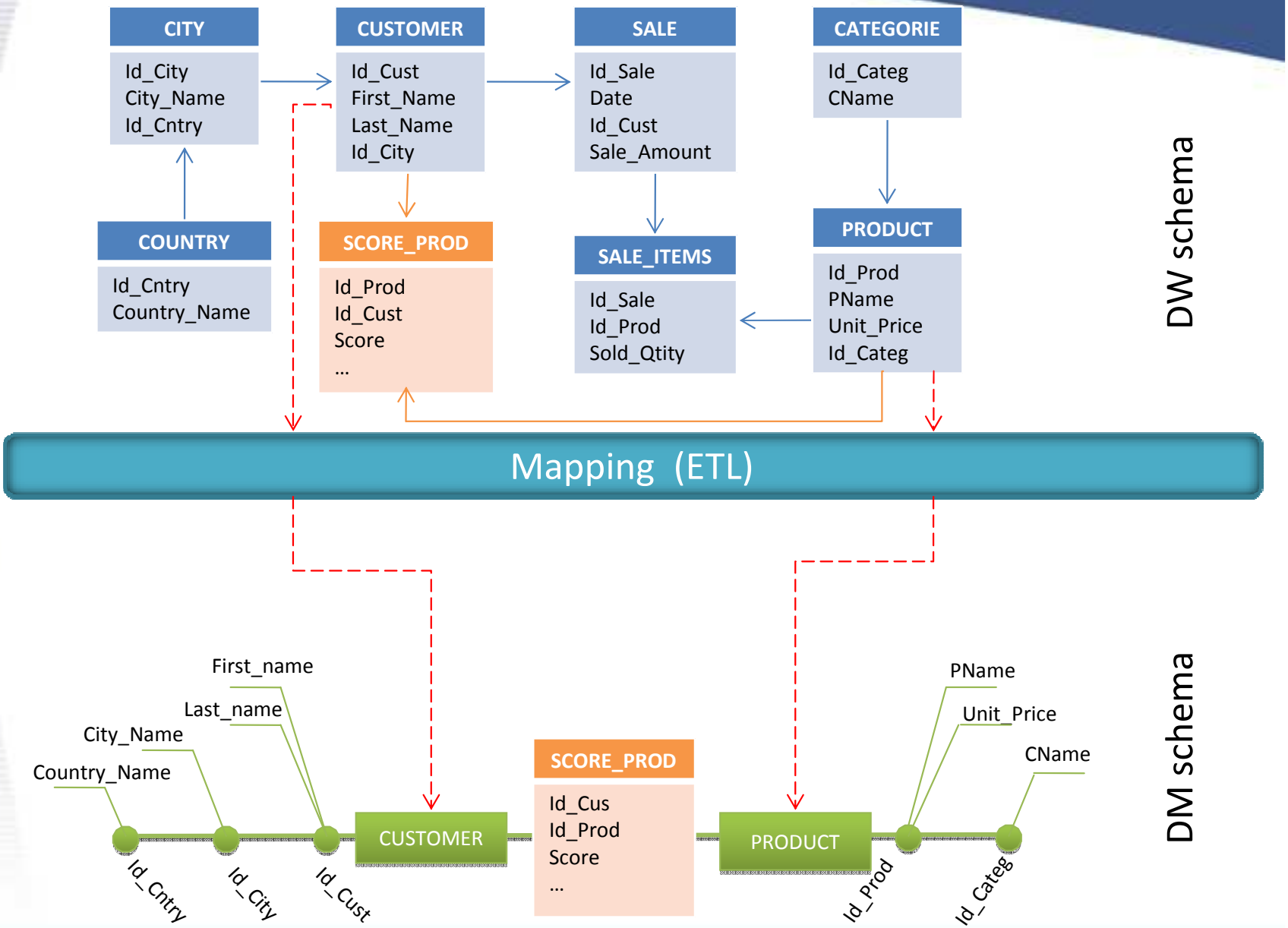
Transformation Rules

- Contexte and motivation
- Related Works
- Approach for Propagating DW Evolution
- ***Transformation Rules***
- Conclusion



Transformation Rules

- Contexte and motivation
- Related Works
- Approach for Propagating DW Evolution
- ***Transformation Rules***
- Conclusion

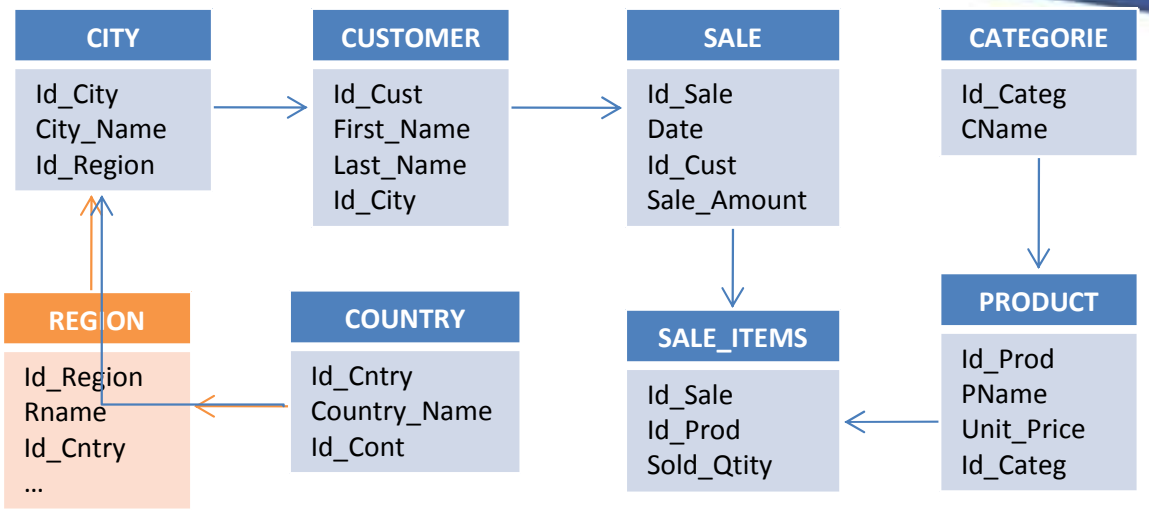


Transformation Rules

- Case (4) : *T inserts a new hierarchy level*
  - *Condition*
    - *If table T is referenced by a DW table T_1 which loads a dimension D of DM*
    - *T references a table T_2 which loads the same dimension D*
  - *Result*
    - *T can feed a new hierarchy level in D*
    - *The identifier of T becomes a parameter*
    - textual attributes of *T are candidate weak attributes for the added level*

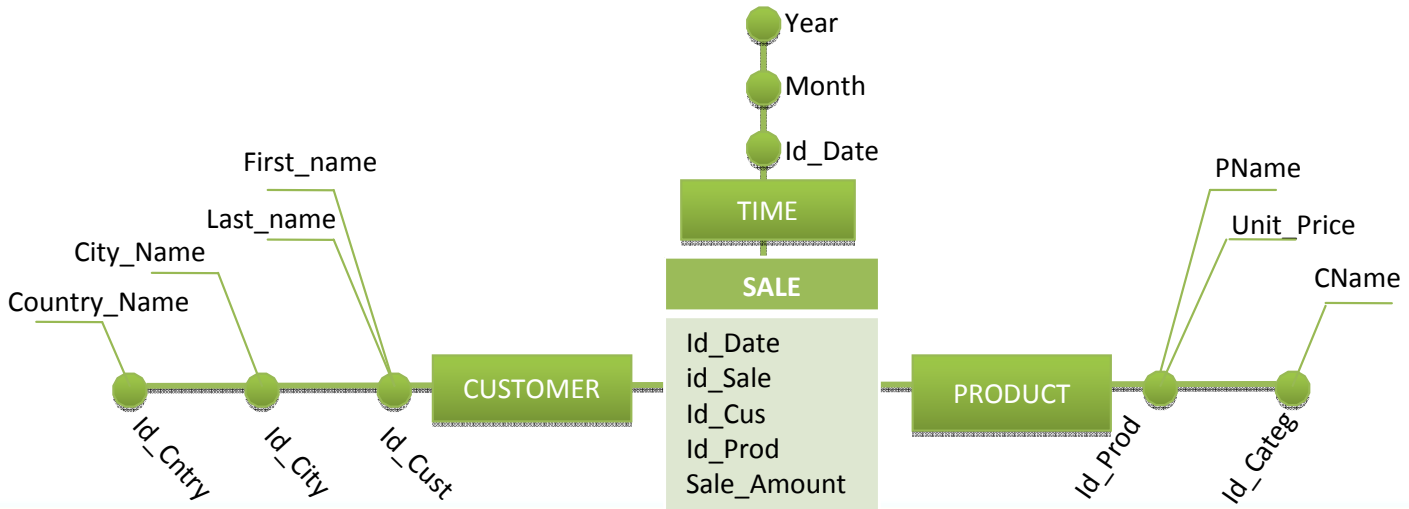
Transformation Rules

- Contexte and motivation
- Related Works
- Approach for Propagating DW Evolution
- *Transformation Rules*
- Conclusion



DW schema

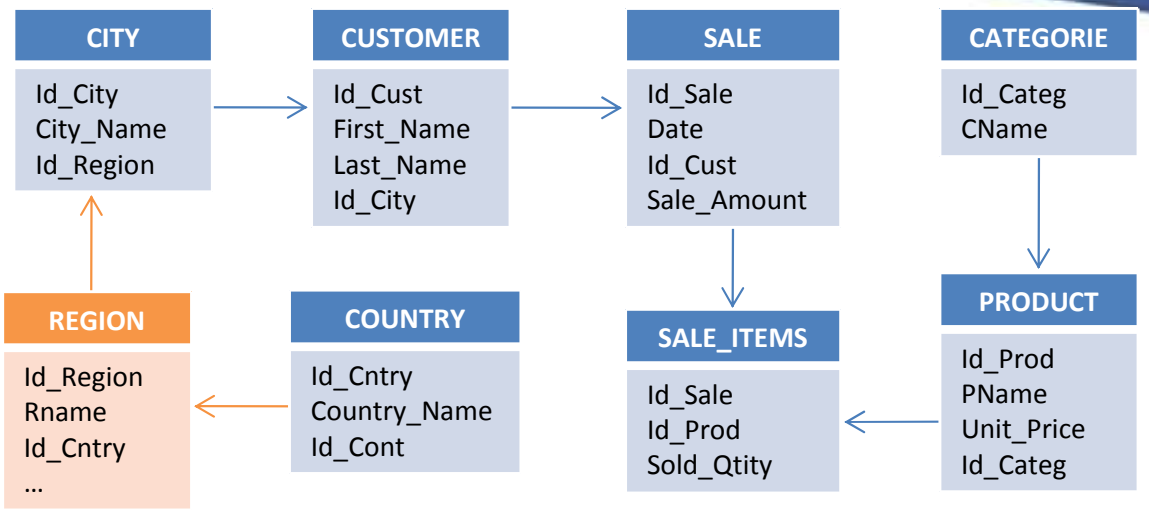
Mapping (ETL)



DM schema

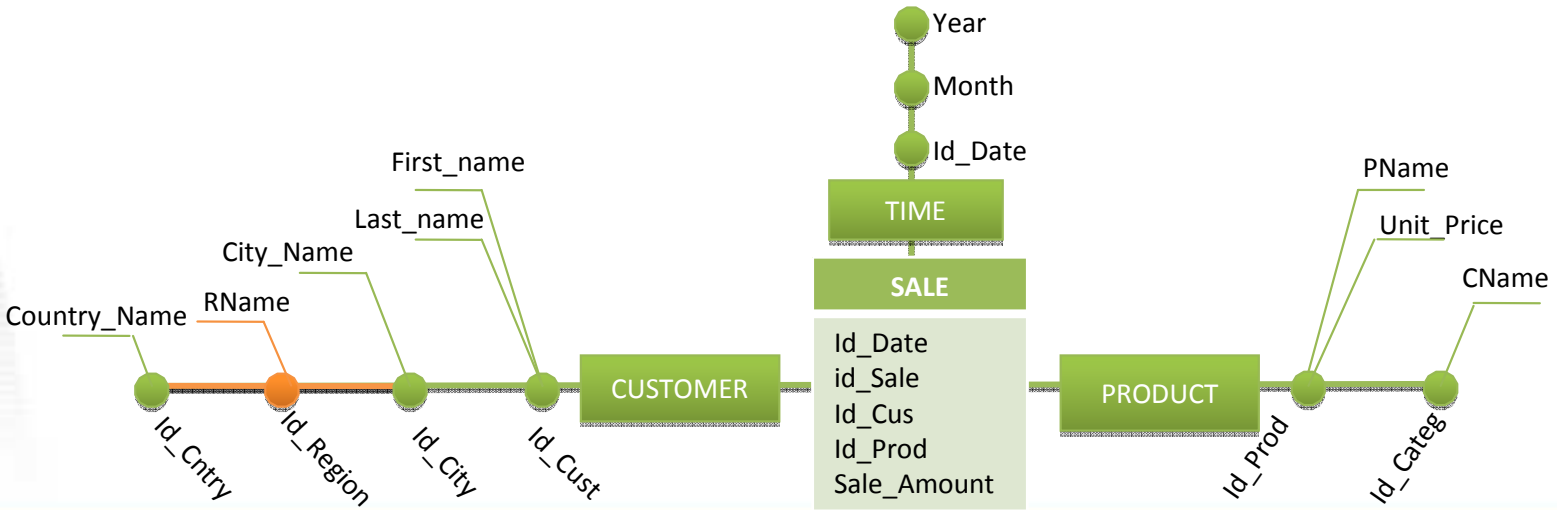
Transformation Rules

- Contexte and motivation
- Related Works
- Approach for Propagating DW Evolution
- *Transformation Rules*
- Conclusion



DW schema

Mapping (ETL)



DM schema

Conclusion

- we studied the impact of the DW schema evolution on DM schemas
  - classified evolution operations into: *basic operations and composite ones.*
  - As a basic operation, we studied the creation of a new table in the relational DW
  - we defined a set of rules and operators.
- Perspectives
  - Currently, we are studying the remaining evolution operations on the DW and their effects on DMs
  - We are also developing a software prototype to apply these transformation rules in the MDA context

Thank you for your attention