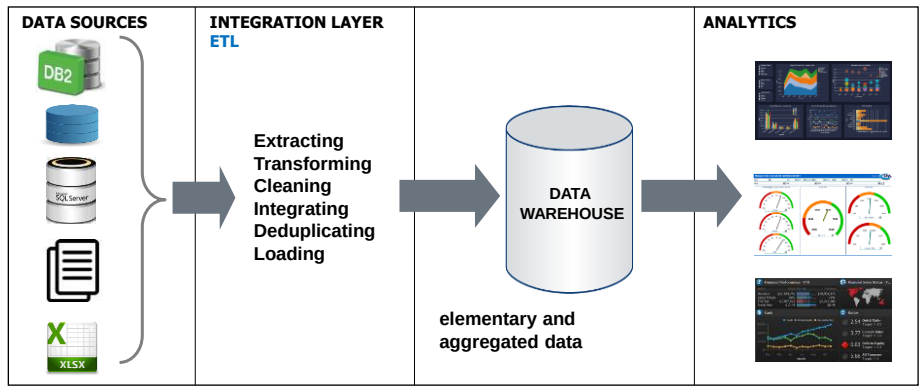
POZNAN UNIVERSITY OF TECHNOLOGY

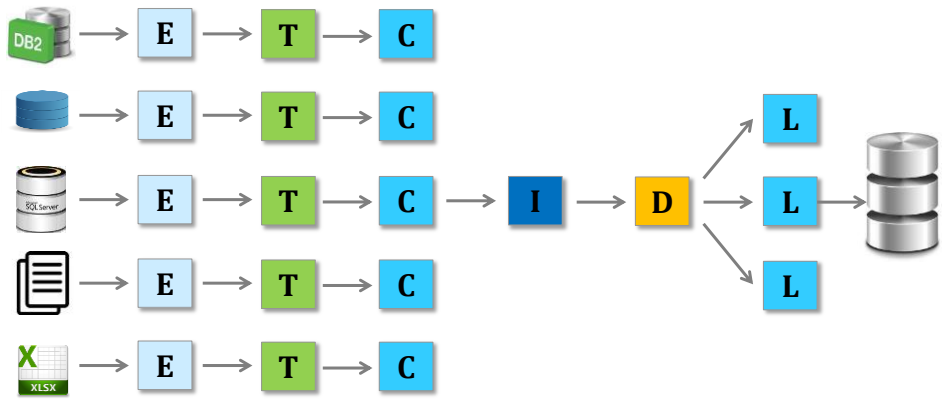
Traditional Data Warehouse Architectures

Robert Wrembel
Poznan University of Technology
Institute of Computing Science
Poznań, Poland
Robert.Wrembel@cs.put.poznan.pl
www.cs.put.poznan.pl/rwrembel

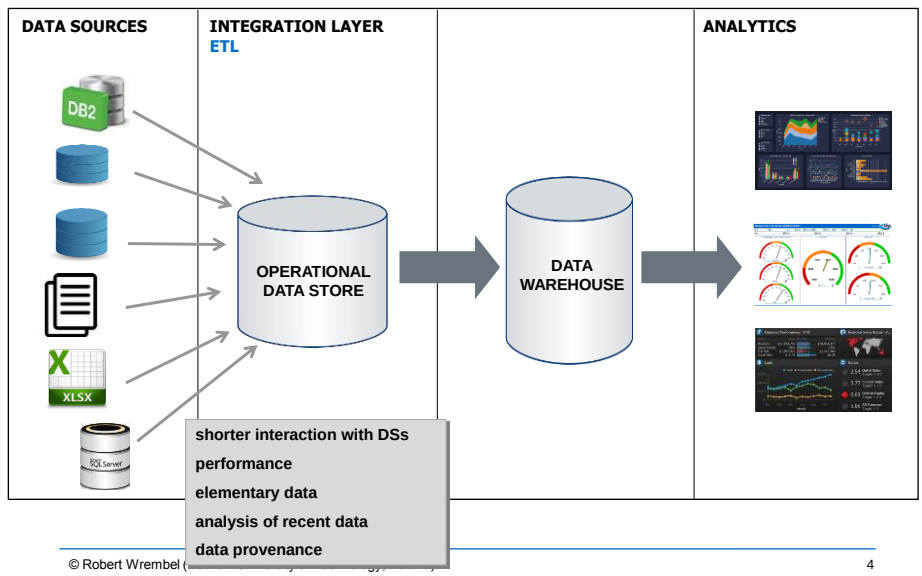
DW Architecture 1 (basic)



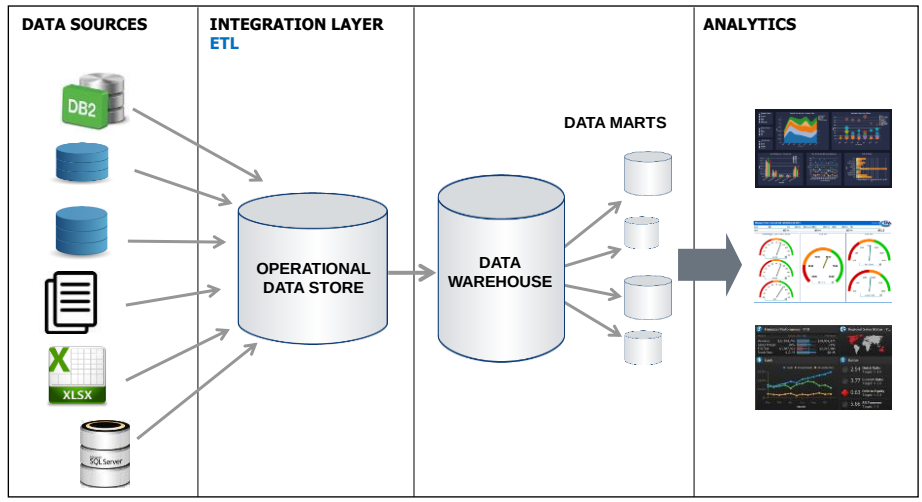
ETL workflow



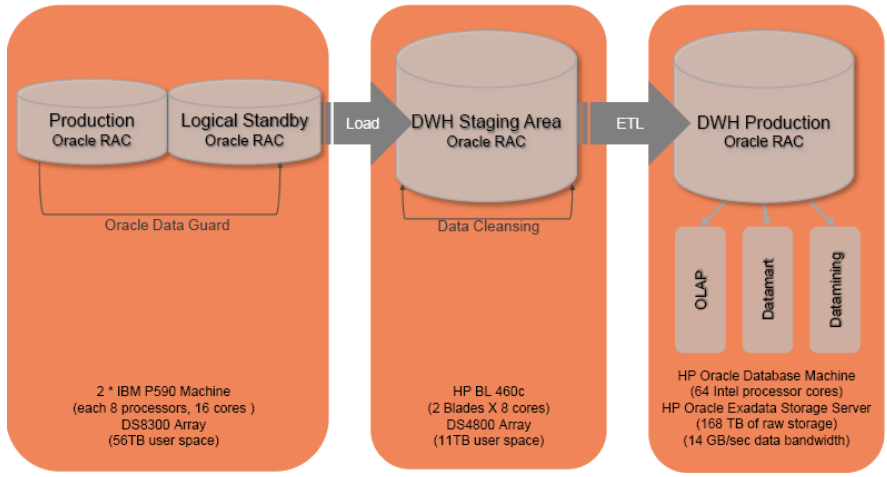
DW Architecture 2



DW Architecture 3

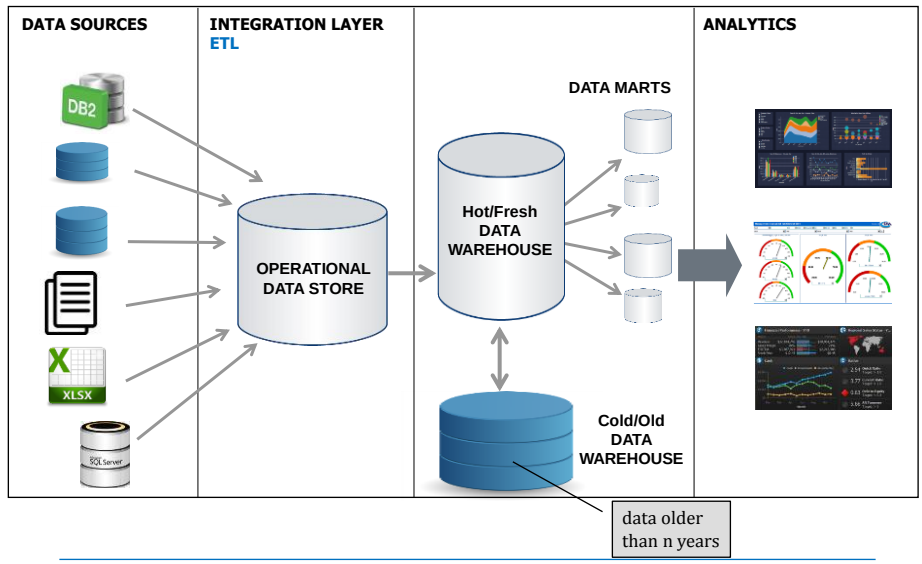


Allegro DW



C. Maar, R. Kudliński: Allegro on the way from XLS based controlling to a modern BI environment. National conference on Data Warehousing and Business Intelligence, Warsaw, 2008

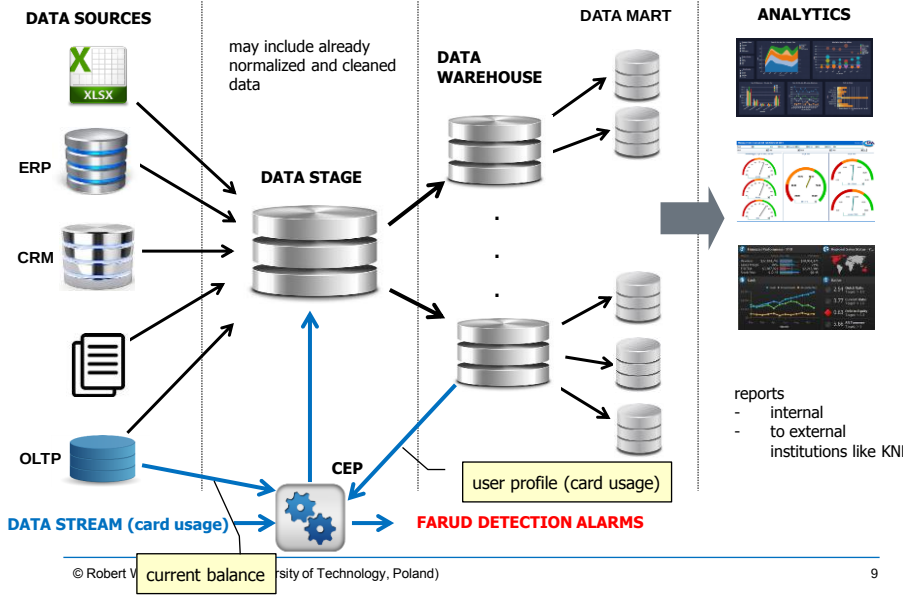
DW Architecture 4



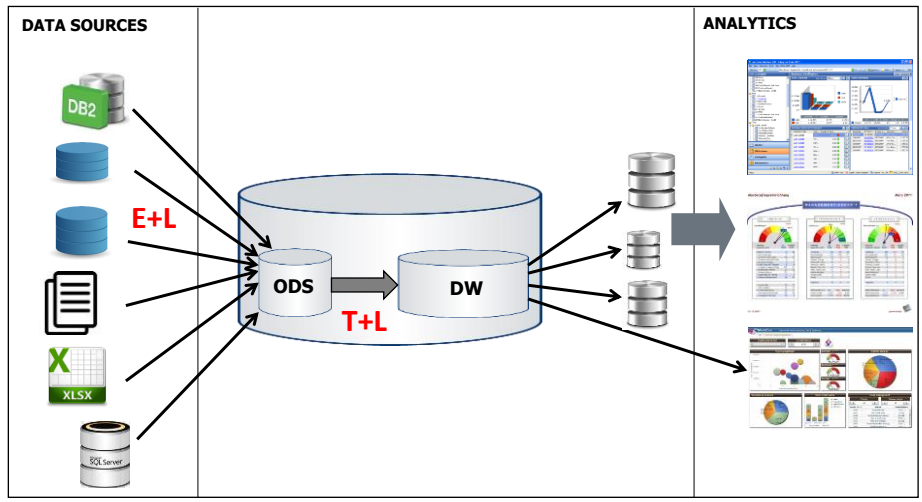
Large DW Architectures

- # data sources: 100 - 200
- Fact table: nn * 10^9 rows
- Fact table: n TB
- Multiple relational DWs in an organization
  - DW size: nn TB
- Multiple data marts
- n * 10^3 to nn * 10^3 ETL workflows

DW in Bank



DW Architecture 5: ELT (ELTL)



ELT Architecture

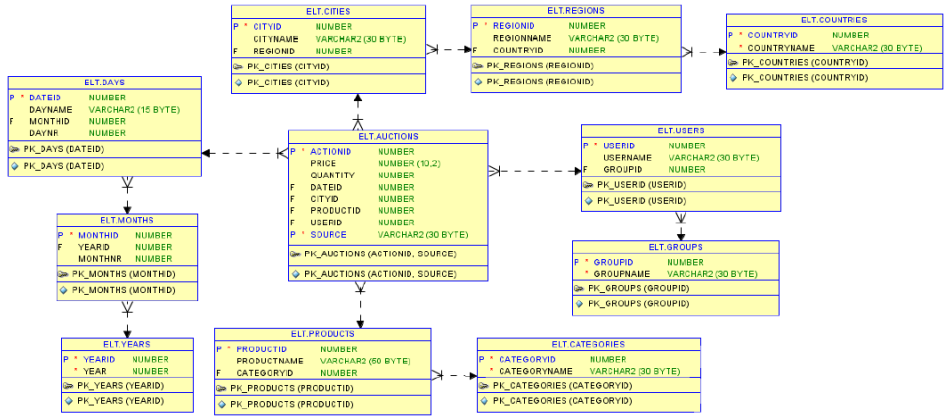
- Performance
  - data stored in a DB ⇨ processing by means of: SQL, PL/SQL, SQL PL, Transact SQL
  - data moved in DB buffer ⇨ native environment
  - advanced query optimization offered by DBMS
  - single server for ELT and HD ⇨ heavier workload
- Functionality
  - data provenance
  - drill through

ETL vs. ELT (experiment 1)

- Data sources
  - topic: Internet auctions
  - storage:
    - Oracle11g (Object-Relational model)
    - MySQL
    - PostgreSQL
    - XML
- Data warehouse: Oracle11g

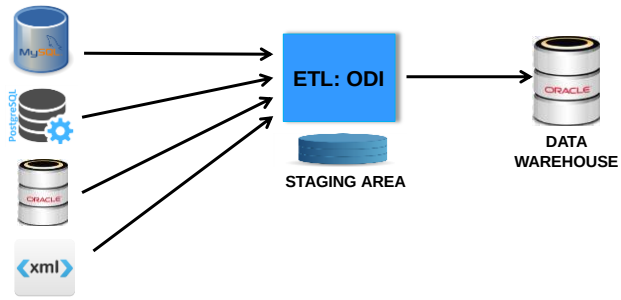
ETL vs. ELT (experiment 1)

- DW schema



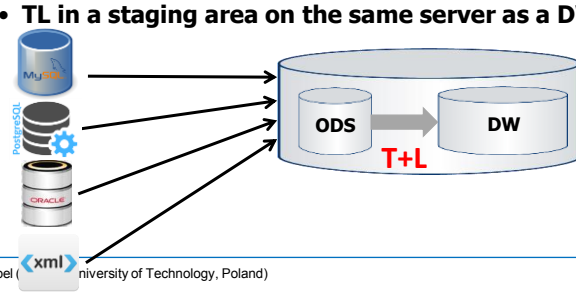
ETL vs. ELT (experiment 1)

- Transformations of data for:
  - dimensions
  - fact table
- Tools and architectures
  - ETL ⇨ Oracle Data Integrator (ODI)
    - ETL in a staging area on a separate server

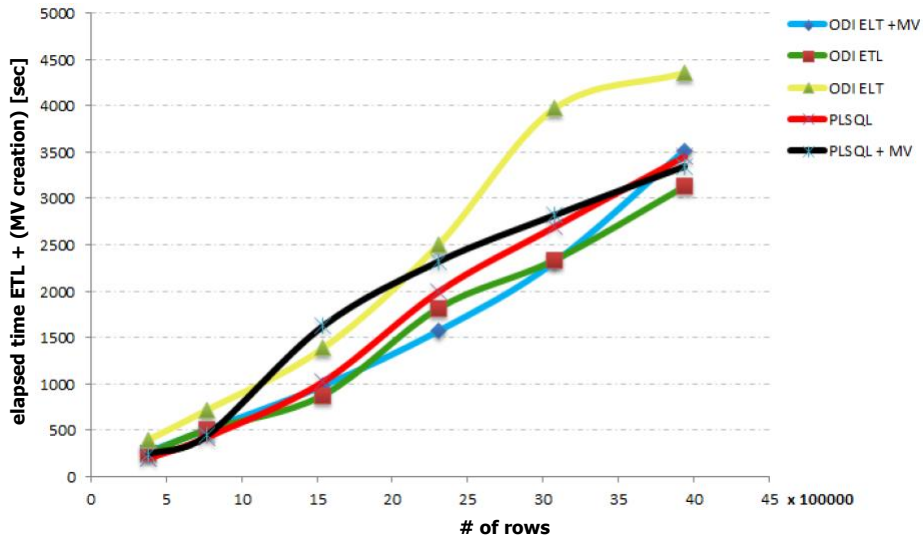


ETL vs. ELT (experiment 1)

- Tools and architectures
  - ELT ⇨ ODI
    - TL in a staging area on the same server as a DW
  - ELT ⇨ ODI + materialized views (MVs)
    - TL in a staging area on the same server as a DW
  - ELT ⇨ stored packages (SPs)
    - TL in a staging area on the same server as a DW
  - ELT ⇨ SPs + MVs
    - TL in a staging area on the same server as a DW

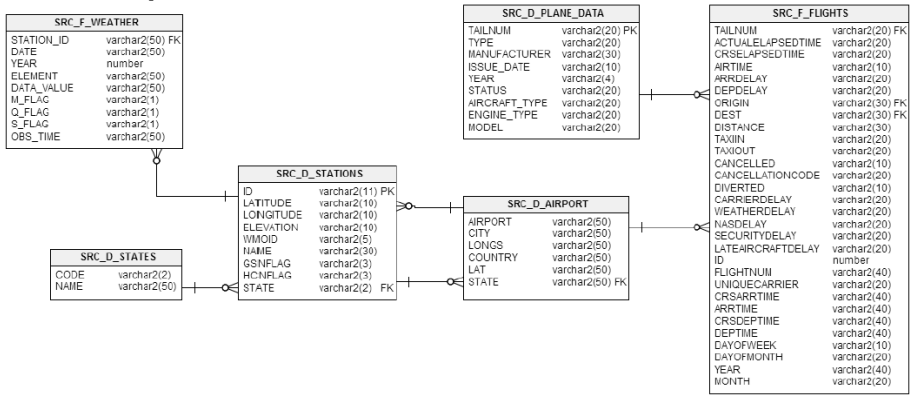


ETL vs. ELT (experiment 1)



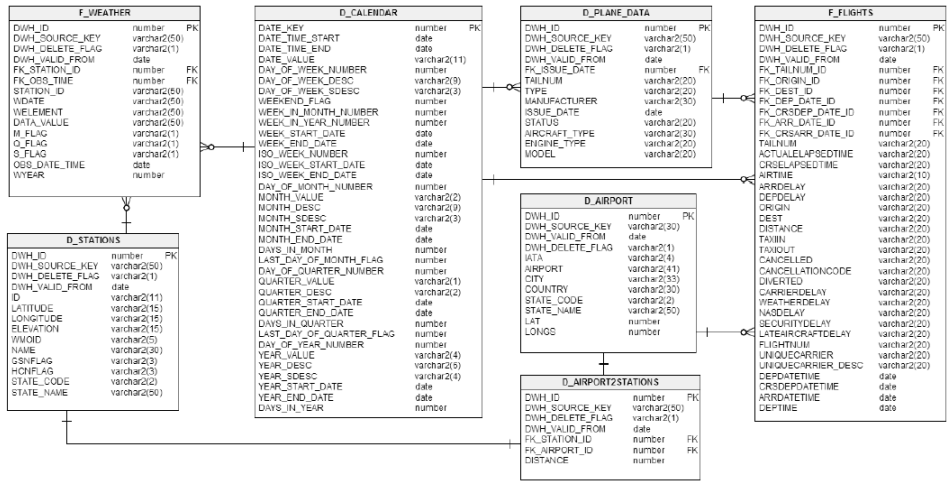
ETL vs. ELT (experiment 2)

- Data source
  - flight and weather data in the US, from 1986 until 2008
  - 6 tables in Oracle11g
- Data warehouse: Oracle11g
- ETL/ELT: Informatica



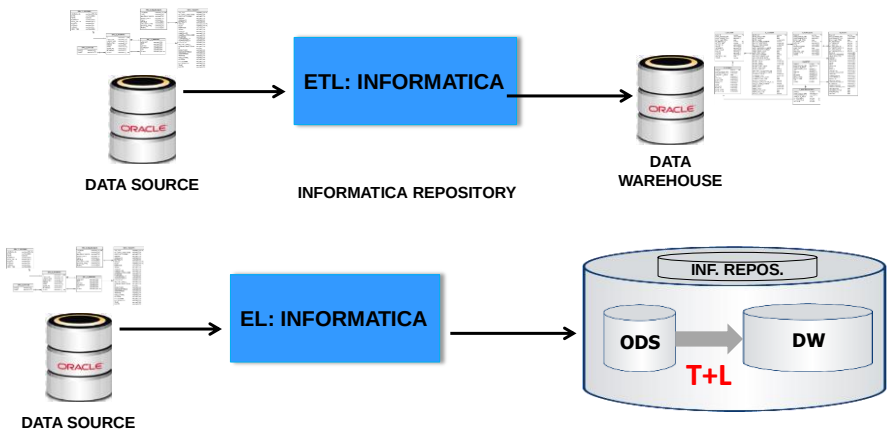
ETL vs. ELT (experiment 2)

- DW schema



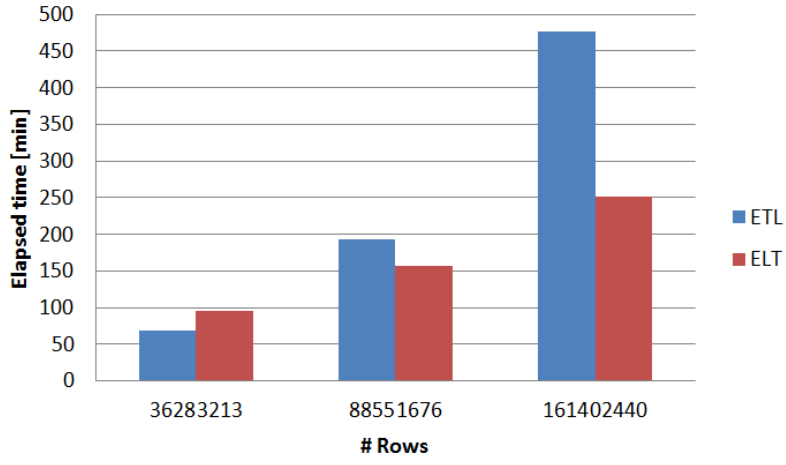
ETL vs. ELT (experiment 2)

- ETL ⇨ Informatica
- ELT ⇨ Informatica (load), DB views (transform)



ETL vs. ELT (experiment 2)

- ETL ⇨ Informatica
- ELT ⇨ Informatica (load), DB views (transform)



Gartner Report: DW servers