



Co-funded by the
Erasmus+ Programme
of the European Union



Pathway in Enterprise Systems Engineering (PENS)

Digital Business integration

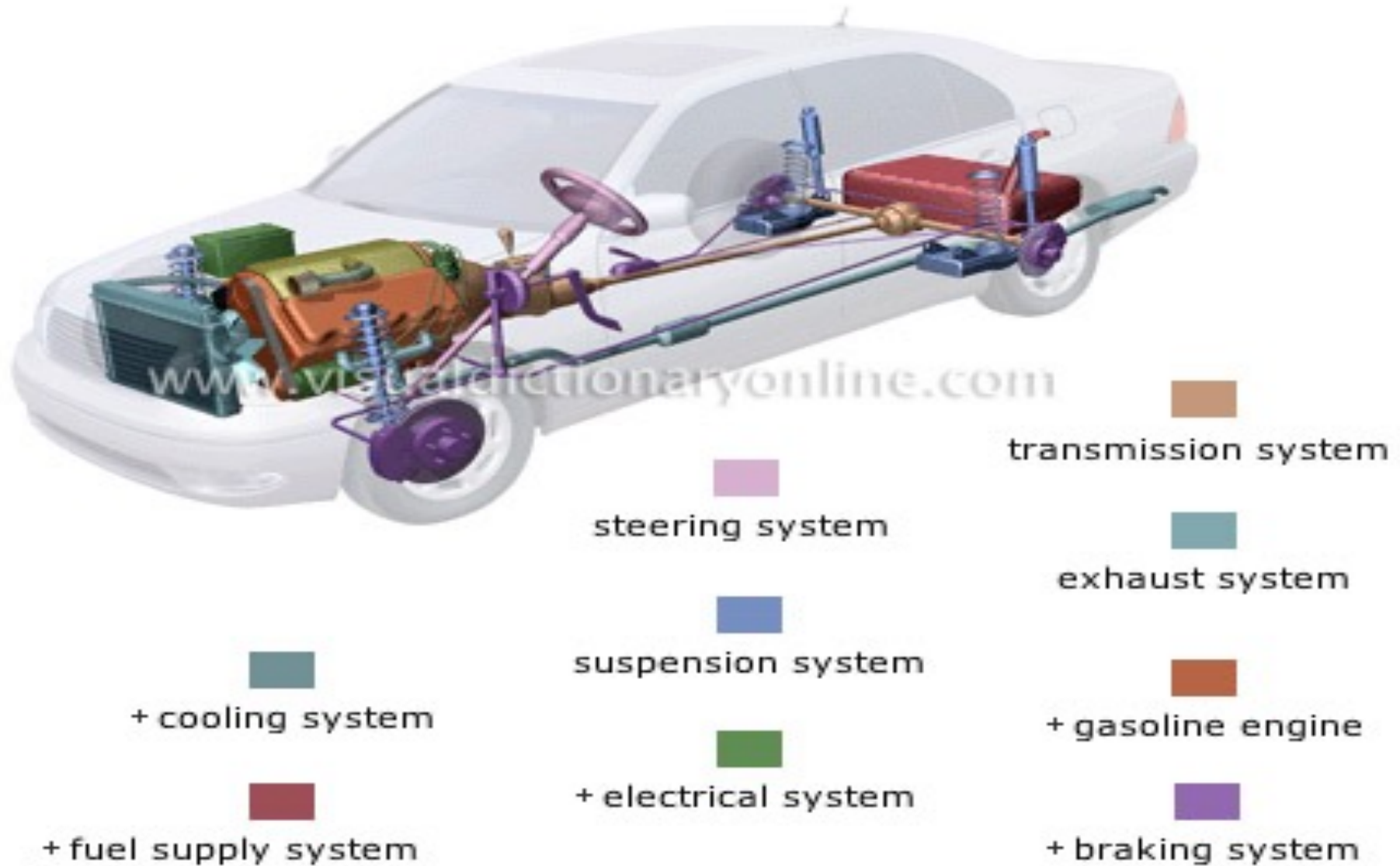
UAH Jul 2022

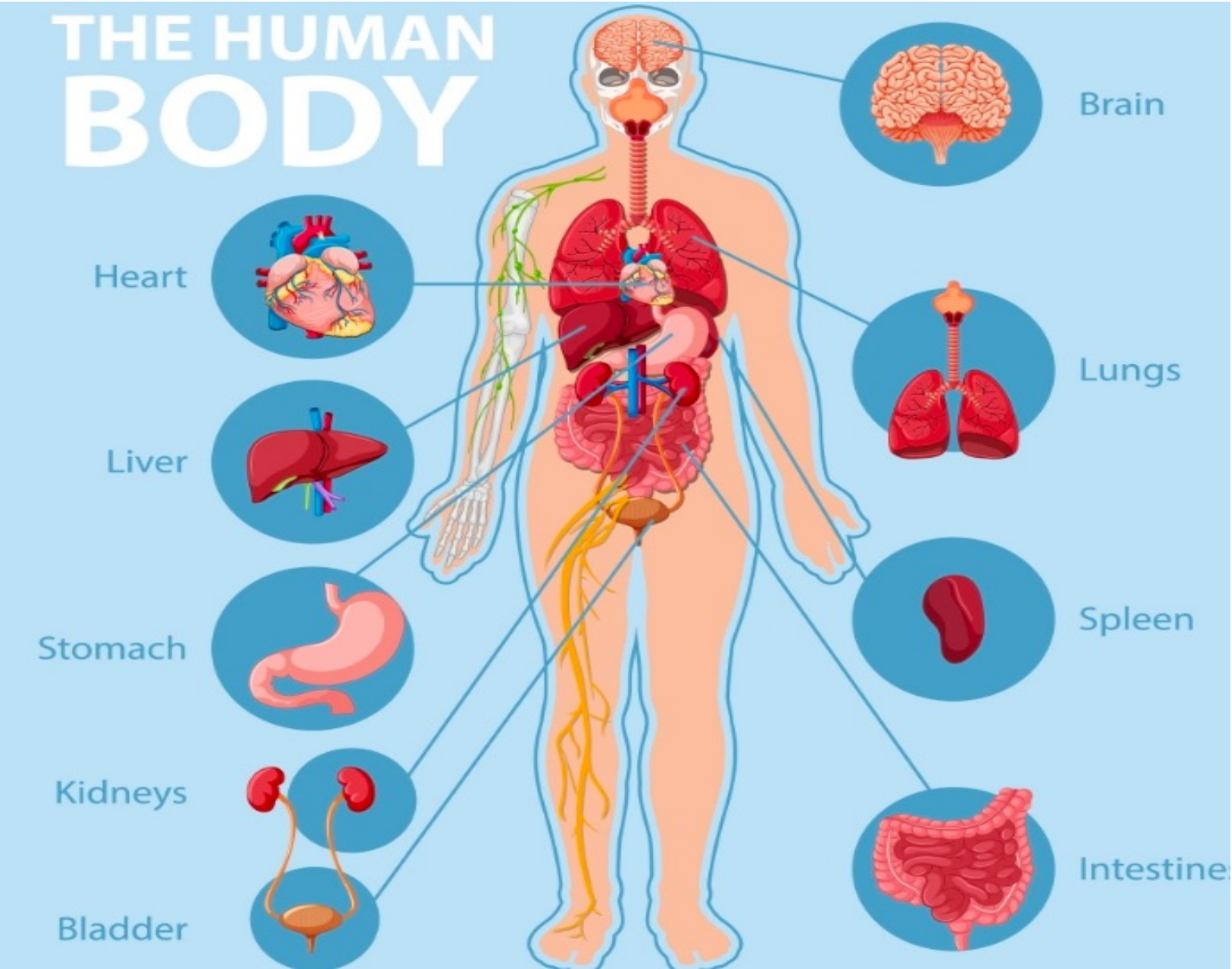
PENS

Pathway in Enterprise Systems Engineering



automobile systems: main parts







Integration vs Interoperability



Example : IT Systems Integration



Email System



So why we need Digital Business Integration ?

Data is growing at an exponential pace, creating massive hurdles for companies in every industry

IT systems are becoming more diverse in every aspect (technology, location etc.)

Companies that can effectively transform their data into insights perform better on almost every business metric

Data Silos

Processes and data are kept in separate servers or data centers and are unable to interact with other systems. (Enterprise Systems)

Data centers grow in size with the deployment of each new application, minimizing resource utilization and reducing the overall efficiency of the enterprise.

Bad Effects

- Limit the view of data
- Threaten data integrity
- Waste resources
- Discourage collaborative work

Cloud Computing ?

Advantages

.....

...

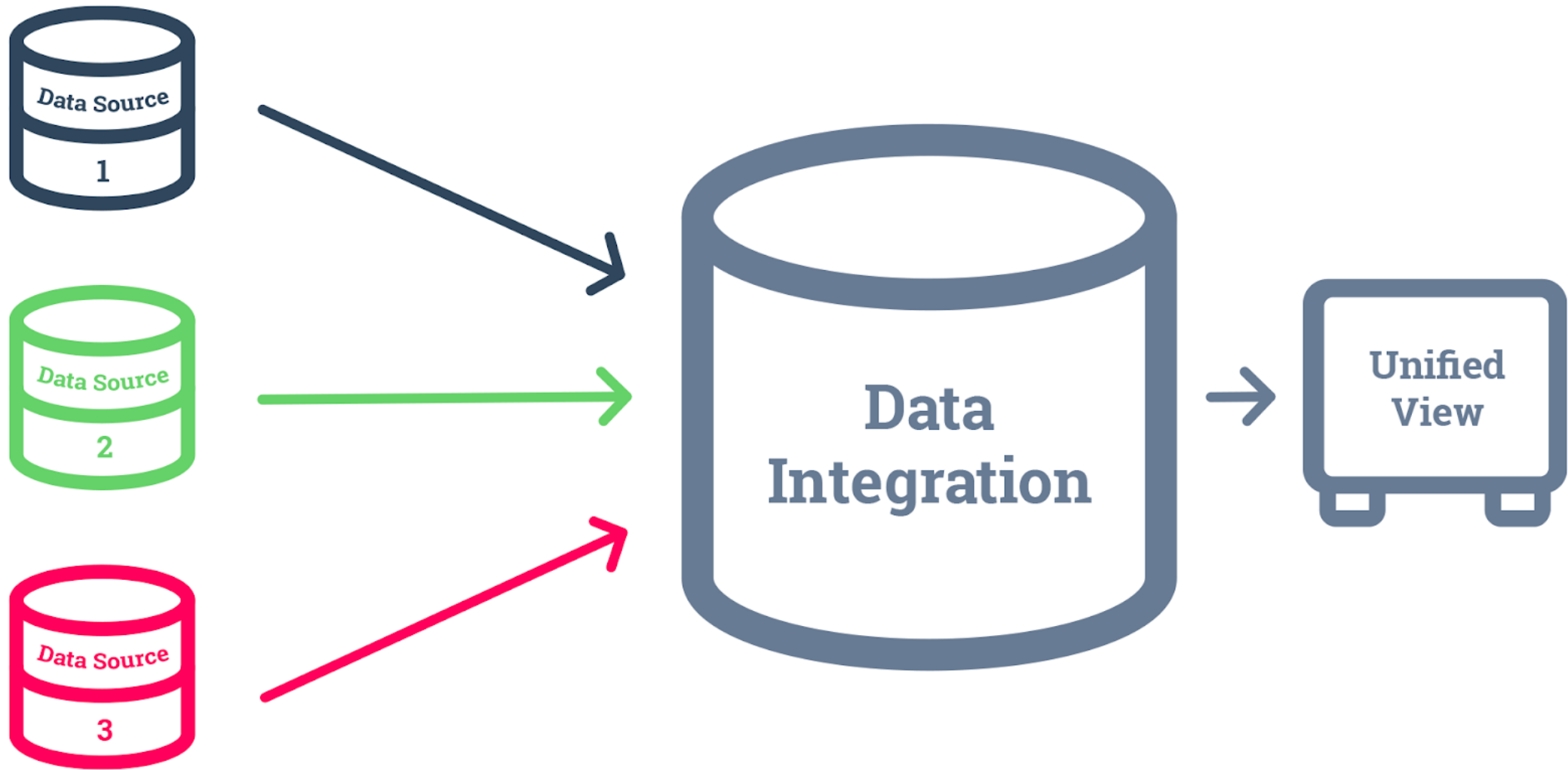
Disadvantages

Data Silos



So does Data Integration help here ?

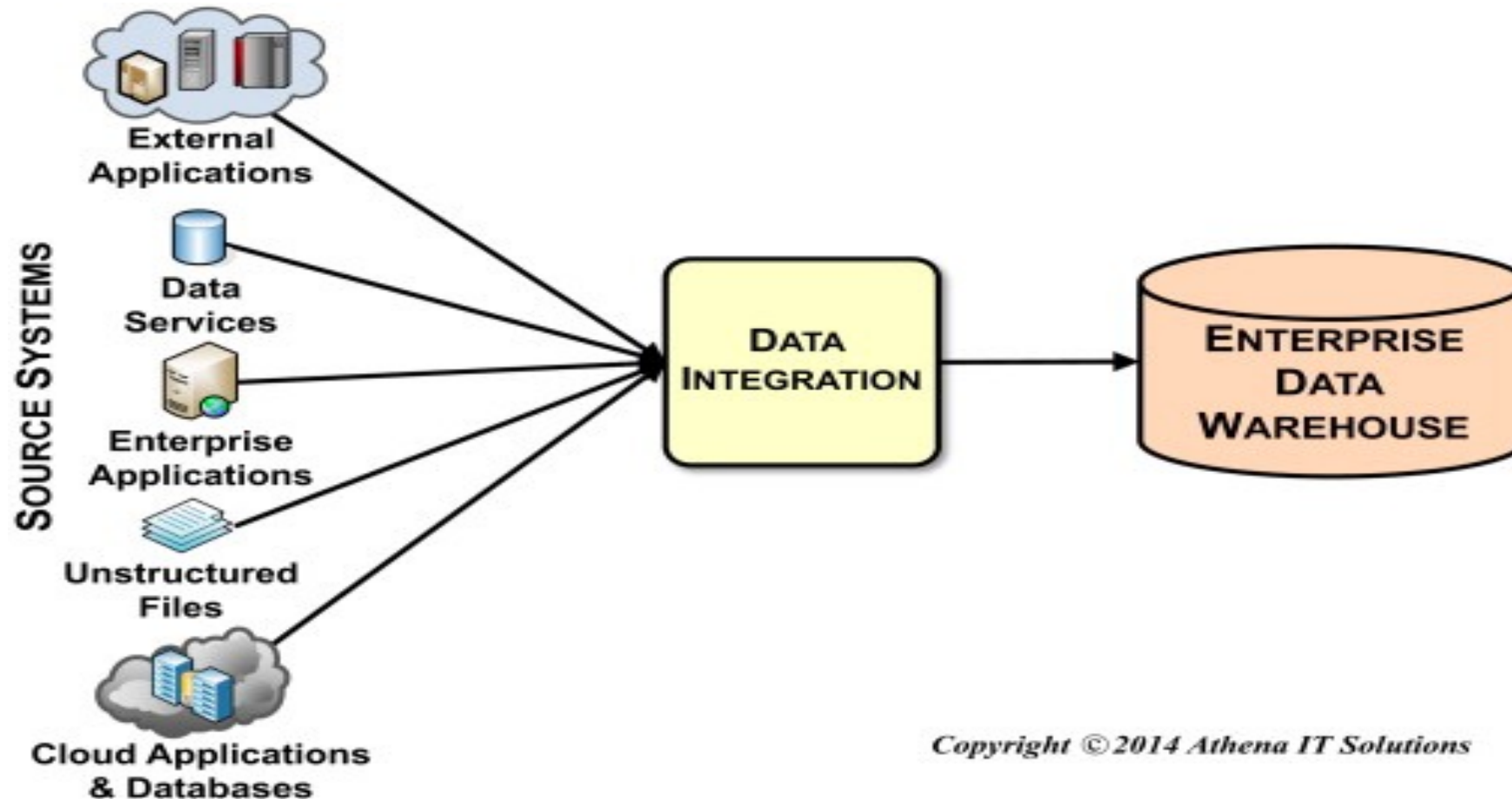




Data Integration Methodologies

- Data transformation (**Aggregate from different resources**)
- Data standardization & cleansing
- **Web services**
- De-duplication
- External data enrichment & validation

Data transformation



Copyright © 2014 Athena IT Solutions

Example

We shall go through an example as follows :

- We need to build/create small databases (multiple ones that store data in local machines)
- Fill these data with transactions

Suggestions For Data sources/Small databases

- H2 database (Java)
- SQLite
- MySQL
- ...

You can select others

Centralized Database

- A Centralized Database will be created
- We shall be using PostgreSQL database
- Small databases (data Sources) have to send transactions to the centralized database and then aggregate data

Sticky Issues

- Transaction IDs
- Pace of sending data (every few min/hours etc) ? Criteria ? Pros and Cons

Why Transaction ID is a sticky issue ?

- Assume **Data source 1** creates Transactions #1 , then # 2, then #3 etc.
- Then **Data source 2** creates Transactions #1,#2 ,#3 etc.
- How to distinguish between these transactions (across space & Time)

Universal IDs

- Universal IDs (UID)
- Unique values that identify transactions /records across space and time
- How they are created ?

Samples

7F000101:01823B97E285:DA3D:7BAE8707

7F000101:01823AA1DA09:CC20:12D3C3A1

5XEkcZX3GOMe8X

They guarantee that each transaction is uniquely identified in the aggregation and transformation process

Practical work to Follow



Requirements

Java

JDBC

Knowledge working with SQL

Practical Exercise 1 : Database Synchronization – Offline- Online Data

This exercise is composed of the following :

- 1- We have a Cloud server running a PostgreSQL a relational database with different sets of tables
- 2- You will be creating a local database in your computer (H2 or mysql etc) and run tests to insert data in the local database .
- 3- You will be periodically sending the data from the Local Database to the online database .
- 4- We shall explain issues in synchronization and avoid problems including using the UUID for table keys etc.