Business School
UNIVERSITY OF COLORADO DENVER

Information Systems Program

Module 1
DBMS Extensions and Example Data Warehouses

Lesson 3: DBMS Extensions

Lesson Objectives

- Review data warehouse characteristics
- Discuss DBMS extensions
- Reflect on importance of DBMS extensions

Data Comparison

- Operational databases
  - Current and some historical
  - Individual
  - Few rows per request
  - Highly volatile
- Data warehouses
  - Historical
  - Individual and summarized
  - Thousands of rows per request
  - Non volatile and refreshed

Database Technology and Deployment Limitations

- Data warehouse technology and deployments
  - Performance limitation
  - Lack of integration
  - Missing data management features

Relational DBMS Dominance for Data Warehouse Processing

- Relational DBMS usage for DWs
  - Relational DBMS dominance
- Lack of scalability of data cube engines
- DW performance improvement
- New DW features

- DBMS Extensions
  - Query language extensions
  - Summary data management
  - Parallel processing
  - Loading and transformation

Query Language Extensions

Motivation

- Limitations of GROUP BY/HAVING clauses
- Tedious and inefficient with SQL and external tools

SELECT statement extensions

- Combine retrieval and analysis
- Subtotal operators
- Analytic processing model and new functions

Summary Data Management

Motivation

- Static environment
- Large number of rows to retrieve per query

Stored queries known as materialized views (MVs)

- Which MVs to store?
- How to combine MVs and user queries?
- When and how to update MVs?

Parallel Processing

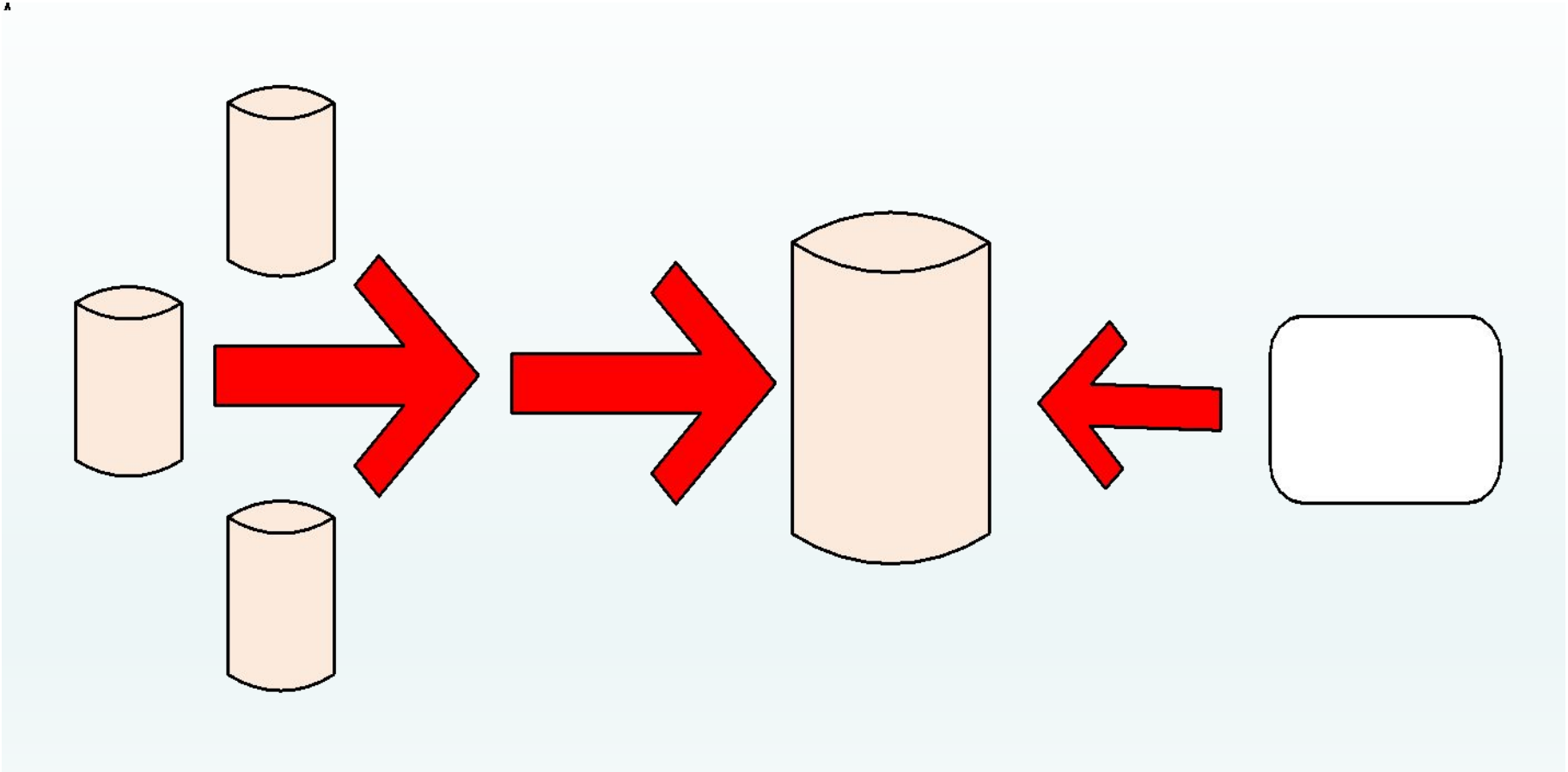
Tasks

- Joins and summary calculations
- Loading
- Data transformations such as parsing and merging

Features

- Transparent
- Scalable
- Commodity components and open source software

DBMS Usage for Loading and Transformation



Summary

- Data warehouse characteristics
- Need for DBMS extensions
- Types of DBMS extensions