

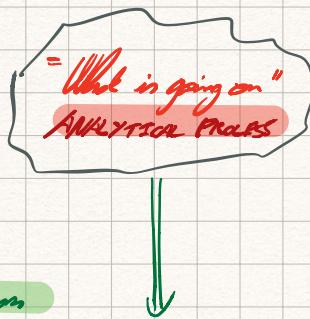
## # Business Perspectives:

Tool to support business activities.

GOT-



Operational Database



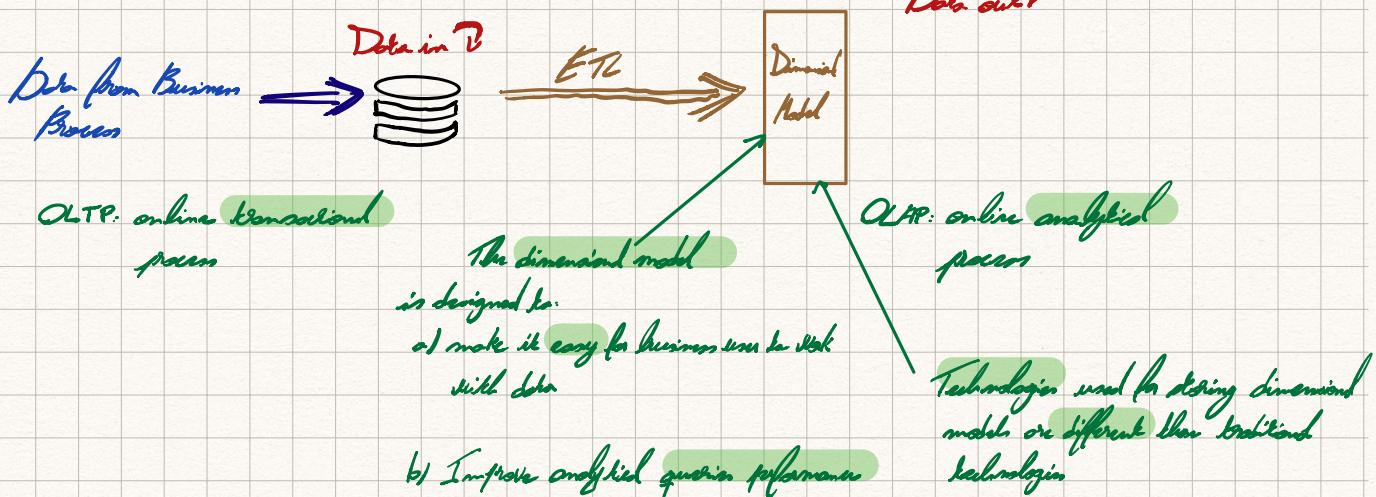
(1)

- Excellent for operations
- No redundancy, high integrity

(2)

- Too slow for analytics; too many joins
- Too hard to understand

## # Data Warehouses: system (including processes, technologies & data representations) that enables us to support analytical processes.



## # Technical Perspective:

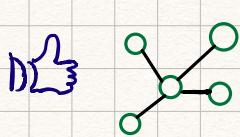
"A data warehouse is a copy of transactional data specifically structured for query and analysis."

"A data warehouse is a subject oriented, integrated, nonvolatile, and time-variant collection of data in support of management's decision"

## # Dimensional Modelling (Recap)

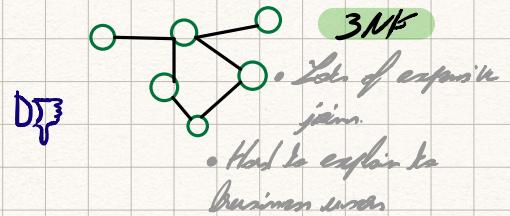
⇒ **Adv.**

- easy to understand
- fast analytical query performance



### Star Schema:

joins with dimensions  
only good for OLAP not  
OLTP



### 3NF

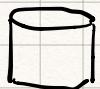
- lots of expensive joins
- Hard to explain to business users

## # End of Dimension dilemma:

Fact & Dimension ⇒ Facts are numeric and additive

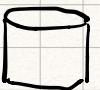
## # DW Architecture: Kimball's Bus Architecture:

Source  
Transactions

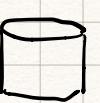


### ETL System:

- Transform from source to target



- Conform dimensions
- No user query support



### Presentation Area:

- Dimensional
- Atomic & Summary data
- Organized by business process
- Uses conformed dimensions



### Applications:

- Ad hoc queries
- Standard reports
- Analytic apps

BACK ROOM



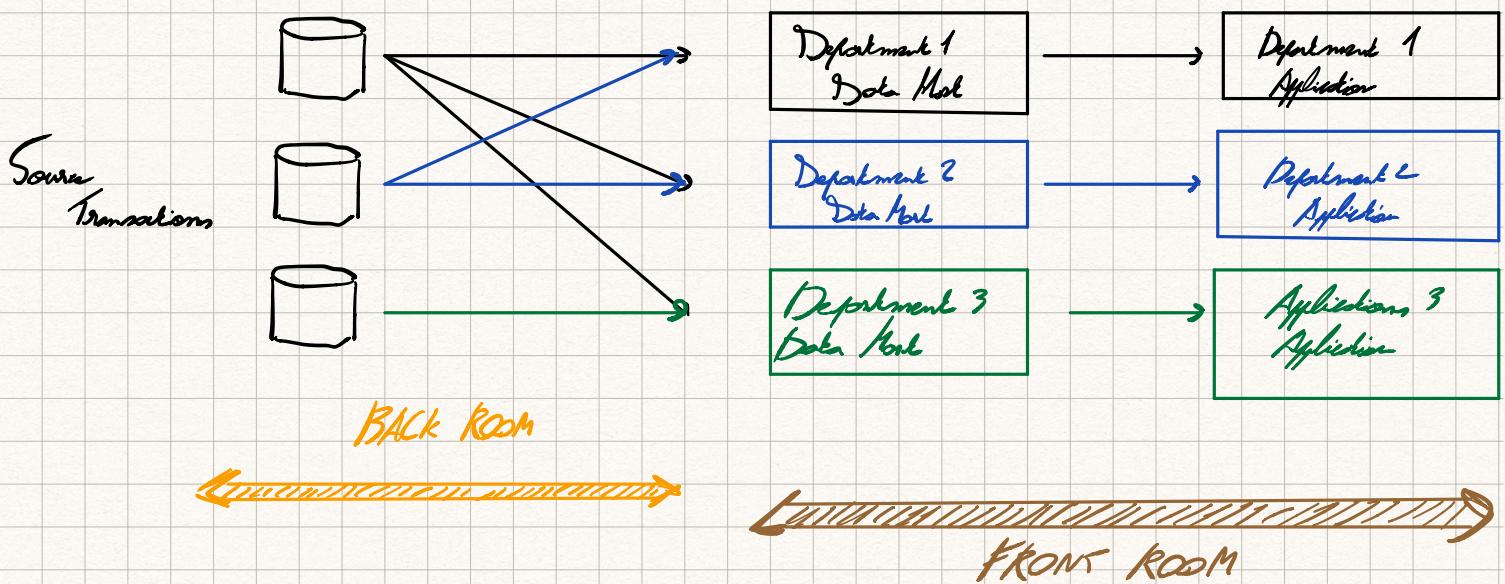
### Design goals:

- Easy of use
- Query performance



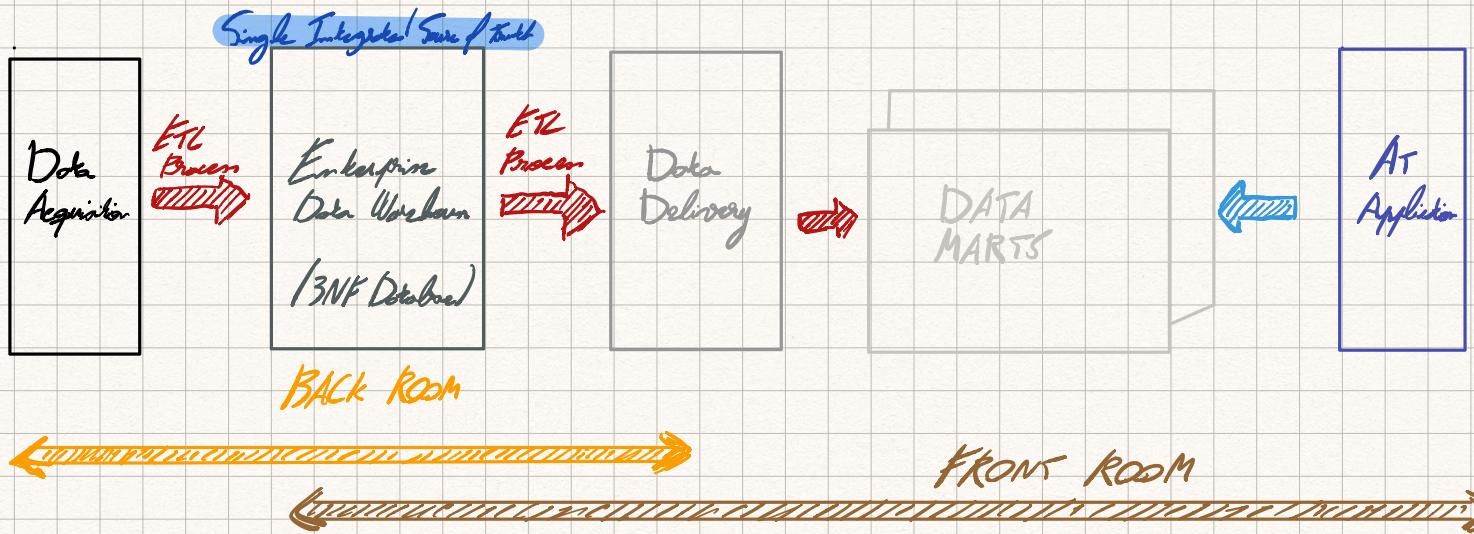
FRONT ROOM

## DWH Architecture: Independent Data Marts



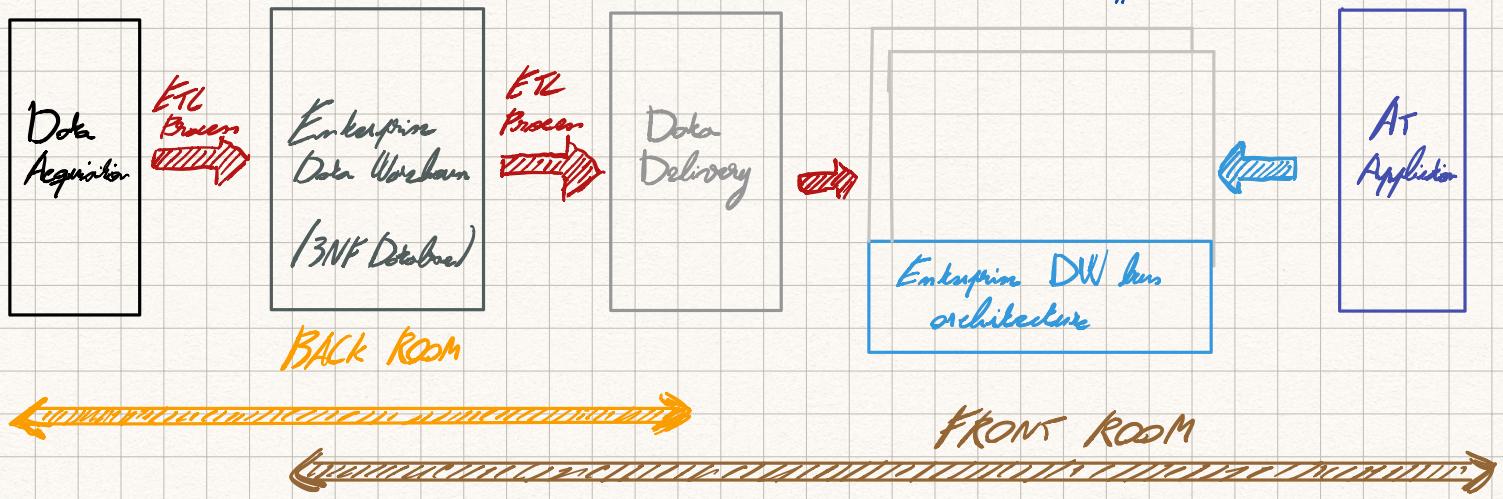
- Department have independent ETL processes & dimensional models.
- Their separate & smaller dimensional models are called "Data Mart".
- Different fact tables for the same event, no conformal dimensions.
- Uncoordinated efforts can lead to inconsistent results.
- Despite awareness of the emergence of this architecture from departmental autonomy, it is generally discouraged.

## Tinmon's Corporate Information Factory (CIF)



## Hybrid Kimball Bus & Inmon C.I.

Source of truth for Application



## # OLAP Cubes: Online Analytical Boxes

- An OLAP cube is an aggregation of a fact metric on a number of dimension: time, brand, store
- Easy to communicate to business users
- Common OLAP operations: Roll-up, drill-down, slice & dice

## # How do we reuse OLAP cubes.

- Approach 1: Pre-aggregates the OLAP cube and stores them on a special purpose non-relational database (MOLAP)
- Approach 2: Compute the OLAP cube on the fly from the existing database where the dimensional model resides (ROLAP)

OLAP

Cube

cities

|      |   |   |   |   |   |   |   |
|------|---|---|---|---|---|---|---|
| time | W | + | * | u | u | u | u |
| a    | 1 | 1 | 1 | " | " | " | " |
| u    | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| u    | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| u    | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| u    | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| u    | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| u    | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Parameters / Attributes

OLAP on revenue

Data (revenue aggregation)