Data Warehouse Architectures

Overview of the Corporate Information Factory and Dimensional Modeling
Rosendo Abellera

- President, BIS3
  - Nearly 2 decades software and system development
  - 12 years in DW and BI space
  - 25+ years of data and intelligence/analytics

- Accenture
- Toshiba
- National Security Agency (NSA)
- US Air Force

Data Warehouse Architect:
- Comcast
- John Hancock Financials
- Manulife
- Engelhard (BASF)
- Mercury
- Lawson
- Global Signal
- Diamond.com
- NationsRent

Other Notable Data Projects:
- LexisNexis
- ESPN
- AAA
- Staples
- Boston College
- US Steel
- British Telecom
- Pfizer
- Toyota
- Partech
What is a Data Warehouse?

A data warehouse is a repository of an organization's electronically stored data designed to facilitate reporting and analysis.

- Subject-oriented
- Non-volatile
- Integrated
- Time-variant

Prevalent Data Warehousing Terms

Enterprise Data Warehouse
Bill Inmon

Corporate Information Factory

Data Mart
Operational Data Store

Snowflake
3rd Normal Form

Dimensional Modeling
Star Schema

Ralph Kimball
Slowly Changing Dimensions
Which Approach Is Each DW Term Most Associated With?

Corporate Information Factory

- 3rd Normal Form
- Bill Inmon
- Data Mart
- Enterprise Data Warehouse
- Hub and Spoke
- Operational Data Store
- Ralph Kimball
- Slowly Changing Dimensions
- Snowflake
- Star Schema

Dimensional Modeling
Overview of Two DW Approaches

- **Corporate Information Factory**
  1. Top down
  2. Data normalized to 3rd Normal Form
  3. Enterprise data warehouse spawns data marts

- **Dimensional Modeling**
  1. Bottom up
  2. Data denormalized to form star schema
  3. Data marts conform to develop the enterprise data warehouse
Corporate Information Factory

- **Focus**
  - Single repository of enterprise data
  - Framework for Decision Support Systems (DSS)

- **Specifics**
  - Create specific structures for distinct purpose
  - Model data in 3^{rd} Normal Form
  - As a Hub and Spoke Approach, create data marts as subsets of data warehouse as needed
Dimensional Modeling

- **Focus**
  - Business Process Oriented
  - User Understandability
  - Performance

- **Specifics**
  - Declare the level of granularity
  - Develop conformed dimensions
  - Identify metrics and measurements
  - Data Bus Matrix
Kimball’s Data Warehouse
Today’s DW Landscape

- Popularity of Dimensional Modeling
  - Adopted by ETL companies (e.g., Informatica)
  - Evident in the strategies of mainstream BI tools such as Cognos, Microsoft’s Analysis Services, OBIEE, etc.

- Shift in focus to Enterprise Architecture
  - SOA and Master Data Management rethinking
  - Consideration for Other Structures
  - Combination of Inmon and Kimball Approaches
Kimball & Inmon Combined

- Holistic Approach to Enterprise Data
- Data Integrity and Cleansing
- DM in Semantic Layer of BI Tools
- ODS can be used as a great data source for the Data Warehouse
Which Approach Is Each DW Term Most Associated With?

Corporate Information Factory

- 3rd Normal Form
- Bill Inmon
- Data Mart
- Enterprise Data Warehouse
- Hub and Spoke
- Operational Data Store
- Ralph Kimball
- Slowly Changing Dimensions
- Snowflake
- Star Schema

Dimensional Modeling
A Data Warehouse By Any Other Name…

- Corporate Info Factory
  - 3rd Normal Form
  - Bill Inmon
  - Hub and Spoke
  - Operational Data Store
  - Enterprise Data Warehouse

- Dimensional Modeling
  - Data Mart
  - Ralph Kimball
  - Slowly Changing Dimensions
  - Snowflake
  - Star Schema
  - Enterprise Data Warehouse
Thank you!

Rosendo Abellera
  › 508.395.1681
  › ross@bis3.com