Data Warehouse Concepts & Architecture
Topics To Be Discussed:

- Why Do We Need A Data Warehouse?
- The Goal Of A Data Warehouse?
- What Exactly Is A Data Warehouse?
- Comparison Of A Data Warehouse And An Operational Data Store.
- Data Warehouse Trends.
Why Do We Need A Data Warehouse?

We Can Only See - What We Can See!
Why Do We Need A Data Warehouse?

BETTER!
FASTER!
CHEAPER!

FUNCTIONALLY COMPLETE!
Data Warehouse Concepts

Data Warehouse Development Perspective

Data Driven Vs. Function Driven

Data

O/P

A/P

DSS

EIS

Order Processing

Data
Data Warehouse Concepts

What Do We Need To Do?

Use Operational Legacy Systems’ Data: To Build Operational Data Store, That Integrate Into Corporate Data Warehouse, That Spin-off Data Marts.

Some May Tell You To Develop These In Reverse!
Our Goal for A Data Warehouse ?

• Collect **Data**-Scrub, Integrate & Make It Accessible

• Provide **Information** - For Our Businesses

• Start Managing **Knowledge**

• So Our Business Partners Will Gain **Wisdom** !
Data Warehouse Concepts

Data Warehouse Definition

A Data Warehouse is a structured repository of historic data.

It is developed in an evolutionary process by integrating data from non-integrated legacy systems.

It is usually:

- Subject Oriented
- Integrated
- Time Variant
- Non-volatile
Data Warehouse Concepts

Subject Oriented

Data is Integrated and Loaded by Subject

- Cust
  - 2001
- Prod
  - 2003
- O/P
  - 2002
- A/R
  - 2002

D/W Data
Data Warehouse Concepts

Time Variant

Operational System

- View of The Business Today
- Operational Time Frame
- Key Need Not Have Date

Data Warehouse

- Designated Time Frame (3 - 10 Years)
- One Snapshot Per Cycle
- Key Includes Date
# Data Warehouse Concepts

## Integrated

<table>
<thead>
<tr>
<th>Operational Systems</th>
<th>Order ID</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order Processing</td>
<td>10</td>
<td>D/W</td>
</tr>
<tr>
<td>Accounts Receivable</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Product Management</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>HR System</td>
<td></td>
<td>M/F</td>
</tr>
<tr>
<td>Payroll</td>
<td></td>
<td>1/2</td>
</tr>
<tr>
<td>Product Management</td>
<td></td>
<td>0/1</td>
</tr>
</tbody>
</table>
Data Warehouse Concepts

Non-Volatile

Operational System

- “CRUD” Actions
  - Create
  - Update

Data Warehouse

- No Data Update
  - Load
  - Read

- Insert
- Read
- Replace
- Delete
- Read
- Read
- Read
Data Warehouse Concepts

Data Warehouse Environment Architecture

Contains Integrated Data From Multiple Legacy Applications

Integration

Criteria

A/P

O/P

Pay

Mktg

HR

A/R

Best System of Record Data

All Or Part Of System of Record Data

Load

Read

Load

Update

Insert

Delete

Replace

Data Warehouse Load Criteria

D/W Load

Read

D/W

Data Mart

Data Mart Loads

Data Mart
Data Warehouse Concepts

Meta Data - Map of Integration

The Data That Provides the “Card Catalogue” Of References For All Data Within The Data Warehouse

- Data Source
- Source Data Structure
- Allowable Domains
- Data Relationships
- System of Record
- D/W Structure
- Definition
- Aliases
## ODS Vs. Data Warehouse

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Operational Data Store</th>
<th>Data Warehouse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data Focused Integration From Transaction Processing Focused Systems</td>
<td>Subject Oriented Integrated Non-Volatile Time Variant</td>
</tr>
<tr>
<td>Age Of The Data:</td>
<td>Current, Near Term (Today, Last Week’s)</td>
<td>Historic (Last Month, Qtrly, Five Years)</td>
</tr>
<tr>
<td>Primary Use:</td>
<td>Day-To-Day Decisions Tactical Reporting Current Operational Results</td>
<td>Long-Term Decisions Strategic Reporting Trend Detection</td>
</tr>
<tr>
<td>Frequency Of Load:</td>
<td>Twice Daily, Daily, Weekly</td>
<td>Weekly, Monthly, Quarterly</td>
</tr>
</tbody>
</table>
### Data Warehouse Concepts

#### Building The Data Warehouse

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Deliverables</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Define Project Scope</td>
<td>• Scope Definition</td>
</tr>
<tr>
<td>• Define Business Reqmts</td>
<td>• Logical Data Model</td>
</tr>
<tr>
<td>• Define System of Record Data</td>
<td>• Physical Database Data Model</td>
</tr>
<tr>
<td>• Define Operational Data Store Reqmts</td>
<td>• Operational Data Store Model</td>
</tr>
<tr>
<td>• Map SOR to ODS</td>
<td>• ODS Map</td>
</tr>
<tr>
<td>• Acquire / Develop Extract Tools</td>
<td>• Extract Tools and Software</td>
</tr>
<tr>
<td>• Extract Data &amp; Load ODS</td>
<td>• Populated ODS</td>
</tr>
</tbody>
</table>
# Data Warehouse Concepts

## Building The Data Warehouse

(Continued)

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Deliverables</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Define D/W Data Reqmts</td>
<td>• Transition Data Model</td>
</tr>
<tr>
<td>• Map ODS to D/W</td>
<td>• D/W Data Integration Map</td>
</tr>
<tr>
<td>• Document Missing Data</td>
<td>• To Do Project List</td>
</tr>
<tr>
<td>• Develop D/W DB Design</td>
<td>• D/W Database Design</td>
</tr>
<tr>
<td>• Extract and Integrate D/W Data</td>
<td>• Integrated D/W Data Extracts</td>
</tr>
<tr>
<td>• Load Data Warehouse</td>
<td>• Initial Data Load</td>
</tr>
<tr>
<td>• Maintain Data Warehouse</td>
<td>• On-going Data Access and Subsequent Loads</td>
</tr>
</tbody>
</table>
Data Warehouse Concepts

Relationship Among Data Warehouse Data Models

- Business Requirements
  - Data Warehouse Physical Model
  - Current Structure
  - Transition Model

- Business Partner
  - Knowledge & Wisdom
  - Validation of Current Data
  - Current Database Physical Model

- Operational Data Store
  - Physical Model
  - Data Load

- Data Whse Requirements
  - Transition Model
  - Strategic Business Requirements
  - Structured Requirements

- Tactical Business Reqmts & Structures
  - Current Database Physical Model
  - Data Warehouse Physical Model

- Operational Data Store
  - Physical Model
  - Current Structure
Sources of Data Warehouse Data

- Archives
  (Historic Data)

- Current Systems
  of Record
  (Recent History)

- Operational Transactions
  (Future Data Source)

Enterprise Data Warehouse
Data Warehouse Concepts

Appropriate Uses of Data Warehouse Data

- Produce Reports For Long Term Trend Analysis
- Produce Reports Aggregating Enterprise Data
- Produce Reports of Multiple Dimensions (Earned revenue by month by product by branch)
Inappropriate Uses of Data Warehouse Data

- Replace Operational Systems
- Replace Operational Systems’ Reports
- Analyze Current Operational Results
Data Warehouse Concepts

Levels of Granularity of Data Warehouse Data

- Atomic (Transaction)
- Lightly Summarized
- Highly Summarized
Data Warehouse Concepts

Options for Viewing Data

- Text

- [Graphs]

- [Graphs]
Data Warehouse Concepts

Next Steps In Data Warehouse Evolution

• Use It - Analyze Data Warehouse Data

• Determine Additional Data Requirements

• Define Sources For Additional Data

• Add New Data (Subject Areas) to Data Warehouse
Data Warehouse Concepts

Future Trends In Data Warehouse

• Increased Data Mining
  Exploration
  Prove Hypothesis

• Increase Competitive Advantage
  (i.e., Identify Cross-selling Opportunities)

• Integration into Supply Chain & e-Business
Data Warehouse Concepts

Summary

A Data Warehouse is a structured repository of historic data.

It is:

- Subject Oriented
- Integrated
- Time Variant
- Non-volatile

It contains:

- Business Specified Data,
  To Answer Business Questions
Data Warehouse Concepts

Questions and Answers