An Overview of Data Warehousing and OLAP Technology

Prabhjot Singh
Contents

- Data Warehousing and OLAP
- Architecture of DW
- Backend Tools and Utilities
- Conceptual Model and Front End Operations
- Database Design Methodology
- Warehouse Servers
- Metadata and Warehouse Management
Data Warehouse

- Integrate Data
- Multiple Heterogeneous Sources
- Support Analytical Reporting
- Decision Making
OLAP

- On-line Analytical Processing
- Use DW data to analyze information
- Data divided into cubes
- Numeric Facts – Measures
- Categorized by Dimensions
## OLAP Cube

### Facts

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of Refusals</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ukraine</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>5000</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>10000</td>
</tr>
<tr>
<td></td>
<td>78</td>
<td>25000</td>
</tr>
<tr>
<td></td>
<td>189</td>
<td>50000</td>
</tr>
<tr>
<td></td>
<td>374</td>
<td>100000</td>
</tr>
<tr>
<td><strong>India</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>81</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>140</td>
<td>5000</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>10000</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>25000</td>
</tr>
<tr>
<td><strong>Kazakhstan</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>98</td>
<td>10000</td>
</tr>
<tr>
<td></td>
<td>178</td>
<td>50000</td>
</tr>
<tr>
<td></td>
<td>389</td>
<td>100000</td>
</tr>
<tr>
<td></td>
<td>435</td>
<td>150000</td>
</tr>
<tr>
<td></td>
<td>782</td>
<td>200000</td>
</tr>
<tr>
<td><strong>Turkmenistan</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>134</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>162</td>
<td>7500</td>
</tr>
<tr>
<td></td>
<td>173</td>
<td>10000</td>
</tr>
<tr>
<td></td>
<td>212</td>
<td>15000</td>
</tr>
<tr>
<td></td>
<td>458</td>
<td>30000</td>
</tr>
<tr>
<td><strong>Armenia</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>5000</td>
</tr>
<tr>
<td></td>
<td>140</td>
<td>10000</td>
</tr>
<tr>
<td></td>
<td>289</td>
<td>20000</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>50000</td>
</tr>
<tr>
<td></td>
<td>530</td>
<td>100000</td>
</tr>
</tbody>
</table>

### Area of Operation

- In the Big City
- More Than 50 KM from the City
- The Mixed
- In the Small City
- Other
Architecture of DW
Backend Tools and Utilities

- Data Cleaning/Transforming
- Loading – Full and Incremental Load
- Refreshing – Updates in Source Data
Conceptual Model and Front End Operations

- Multidimensional data model (Cube)
- Pivoting
- Rollup
- Drill Down
- Slice and Dice
Database Design Methodology

- Star Schema – Single Fact Table and Single Table for each dimension
Snowflake Schema – Refinement of Star, dimension tables are normalized.
<table>
<thead>
<tr>
<th>Basis for comparison</th>
<th>ROLAP</th>
<th>MOLAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage &amp; fetched</td>
<td>Data stored and fetched from main data warehouse.</td>
<td>Data stored and fetched from the MDDBs.</td>
</tr>
<tr>
<td>Data Form</td>
<td>Relational tables.</td>
<td>Large multidimensional array made of data cubes.</td>
</tr>
<tr>
<td>Data volumes</td>
<td>Large data volumes.</td>
<td>Limited summaries data is kept in MDDBs.</td>
</tr>
<tr>
<td>View</td>
<td>Creates a multidimensional view of data dynamically.</td>
<td>Already stores the static multidimensional view of data in MDDBs.</td>
</tr>
<tr>
<td>Access</td>
<td>Slow access.</td>
<td>Faster access.</td>
</tr>
</tbody>
</table>
Metadata and Warehouse Management

- **Types of Metadata:**
  - Administrative Metadata
  - Business Metadata
  - Operational Metadata

- **Metadata Repository**
  - Store and Manage
  - Share among tools
Thank You!