Building data warehouse in Metropolia University of Applied Sciences

Antti Tikka, 12.6.2013
Introduction

Inmon

Centralized DW

Operative databases

Data marts

Kimball

Data marts

Operative databases
1. **Starting with Inmon’s architecture**
2. Problems
3. Kimball’s architecture
Conceptual model of the centralized DW

- Design started in 2006
- Co-operation with consultants and our own specialists in different subject areas
- First version ready in 2008
- 120 different concepts of studying, teaching, HR, organization ja finance
Implemented architecture in 2010

Operative databases:
- Course feedback
- HR
- Studying data
- Strategic planning
- Files

Data marts:
- Course feedback
- Centralized DW
1. Starting with Inmon’s architecture
2. **Problems**
3. Kimball’s architecture
Problems in centralized DW

- Complex queries
- Empty tables and columns
- Old data in some tables
- Major focus in centralized DW and only little in data marts

=> Need for data marts
Problems in ETL-processes

- Two data models and two ETL-processes
- => Lots of work and many errors
- Tracking error: Is it in first or second ETL-process?
1. Starting with Inmon’s architecture
2. Problems
3. Kimball’s architecture
Principle of Kimball’s architecture

- No centralized DW
- DW is just a group of data marts
- Data marts are compatible to each other
- Data warehouse is built one data mart at a time
Star schema

- Core of Kimball’s architecture
- Consist of fact and dimension tables
- Diverse aggregating possibilities
- Simple and efficient queries
- Work well with automatic query generation tools
- Easily transformed into OLAP-cube
Conformed dimensions
# Bus matrix

<table>
<thead>
<tr>
<th>Dara mart / fact table</th>
<th>Fact</th>
<th>Date</th>
<th>Month</th>
<th>Quartal</th>
<th>Season</th>
<th>Year</th>
<th>Implementation</th>
<th>Course</th>
<th>Degree program</th>
<th>Studying area</th>
<th>Employee</th>
<th>Student</th>
<th>Grade</th>
<th>Question</th>
<th>Reason</th>
<th>Cost center</th>
<th>Account</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Studying</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achievement</td>
<td>Number of credits</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence</td>
<td>True / false</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Course feedback</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answer to question</td>
<td>Numerical assessment</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open feedback</td>
<td>Textual assessment</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Finance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounting event</td>
<td>Amount of money</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Budget</td>
<td>Amount of money</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Change of architecture

- Do we have to replace already implemented Centralized DW?
- No, just leave implemented structures as they are
- Continue development using Kimball’s architecture
Data warehouse architecture in 2013

Operatiiviset tietokannat
- Course feedback
- HR
- Studying data
- Strategic planning
- Files
- Accounting

Centralized DW

Datamartit
- Course feedback
- Studying
- Finance

Metropolia

12.6.2013
Antti Tikka
References


Thank you!

www.metropolia.fi
www.facebook.com/MetropoliaAMK
antti.tikka@metropolia.fi