Multidimensional Data Model Extensions to Data Warehouses

Mikael Hagström
SAS Institute
## Different Needs

<table>
<thead>
<tr>
<th>Information Technology</th>
<th>Knowledge Workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>☀ Enable end-users fast access to summarized data</td>
<td>☀ Satisfy need to have data reflect their business vision</td>
</tr>
<tr>
<td>☀ Alignment with data warehouse architecture</td>
<td>☀ Exploratory ‘analysis’, help to identify possible answers</td>
</tr>
<tr>
<td>☀ Accurate and timely data</td>
<td>☀ Access to subsets of detail data</td>
</tr>
<tr>
<td>☀ Build, maintain and perform capabilities</td>
<td>☀ Simple (i.e. spreadsheet-like) access</td>
</tr>
<tr>
<td>☀ Interested in hardware and network demands</td>
<td></td>
</tr>
</tbody>
</table>

What can SAS Institute offer them?
Business Intelligence Architecture

Warehouse Administrator

Extract
Transform
“Scrub”
Add-value

Warehouse Storage Architecture

OLAP

MDDB

Applied Analysis
“Packaged OLAP” Technology

- Proprietary data storage architecture
  - MDDB (Multi Dimensional Data Base)
    - SAS/EIS Class

- Viewers
  - MDDB Enabled SAS/EIS Classes
    - Multidimensional report
    - 3D Business graphs
    - Map
    - Organizational chart
    - Graphical variance report
Data Warehouse Storage Structure

Metadata

MDDB
Hybrid OLAP (HOLAP)

- **Relational OLAP (ROLAP)**
  - Relational Storage
    - highly voluminous data
    - typically flexible

- **Multidimensional OLAP (MOLAP)**
  - Multi Dimensional Storage (MDDB)
    - summarisable data
    - typically fast
SAS MDDB - Features

- Efficient storage, Sparsity control
- Flexible tuning
- Fast summarization
- Automatic statistics
- Fast information retrieval
- Flexible Client/Server deployment
- Reach through to source data
- Multi-platform ("Mobile OLAPs")
- Integrated with
  - Data Warehouse
  - Analysis
  - PC Personal Productivity tools
Multi-Dimensional Data is typically Sparse by its nature

... but clustered.
MDDB - What’s Inside?
MDDB - What’s Inside?

Subtable

Lookup table

NWAY Crossing
MDDB - How does it work?

Request

Reach Through

NWAY Crossing
Size Relationship

Operational Data

MDDB
Sample PC Based MDDB

<table>
<thead>
<tr>
<th>Base Table:</th>
<th>MDDB Class:</th>
</tr>
</thead>
<tbody>
<tr>
<td>◆ 35.5 Million Records</td>
<td>◆ 19.2 Million Potential Cells</td>
</tr>
<tr>
<td>◆ 2.2 % Sparsity</td>
<td>◆ 4.2 Million Input Cells</td>
</tr>
<tr>
<td>◆ 1.5 Gb Storage</td>
<td>◆ 378 Mb Storage</td>
</tr>
<tr>
<td></td>
<td>◆ 65 Mb Minimal Storage</td>
</tr>
</tbody>
</table>

**Characteristics:**

- 5 Dimensions
- 10 Classification Groups
- 4 Prestored Statistics
Remote Data

Exploit Improved RLS

- Higher Transfer Speed
- Less Data Transferred
## Sample Server Based MDDB

### Base Table:
- **429 Million Records**
- **29.8% Sparsity**
- **9 Gb Storage**

### MDDB Class:
- **94 Million Potential Cells**
- **28 Million Input Cells**
- **1.7 Gb Storage**
- **1 Gb Minimal Storage**

### Characteristics:
- **5 Dimensions**
- **11 Classification Groups**
- **4 Prestored Statistics**
MDDB Creation

- Interactively using EIS
- Interactively using SCL
- Batch using Proc MDDB
- Can be created and/or accessed on any host
  - eg: MVS, VM, Unix, Windows, …
- Including Partial Update
Demo
SAS MDDB - Features

- **Efficient** storage, Sparsity control
- **Flexible** tuning
- Fast summarization
- Automatic statistics
- **Fast** information retrieval
- Flexible Client/Server deployment
- Reach through to source data
- Multi-platform ("Mobile OLAPs")

- **Integrated** with
  - Data Warehouse
  - Analysis
  - PC Personal Productivity tools
Business Intelligence Architecture

Warehouse Administrator
- Extract
- Transform
- "Scrub"
- Add-value

Warehouse Storage Architecture

Applied Analysis

OLAP

MDDB
Thank you for your attention

The SAS® System for successful decision making