STATEMENT:

BI warehouses are slowly moving dinosaurs

Wolfgang Lehner
General Observation

• ... from 10.000 feet

  • BI/DWH-environment has to fulfill two basic tasks

    – consolidation
      model and schema integration,
      logical/physical data integration, ...

    – stability
      analytical business processes
      foundation of statistical correctness

...here is an example !!!
**S T A R T R A C K**: data production chain

- **Retailers**
  - 1000s of sources/country
  - independent data delivery
  - fact data (sales, price, turnover)

- **Production, Analysis & Reporting**
  - consolidation process
  - multiple data refinement steps

- **Data Warehouse**
  - IDAS
  - 450,000 different products with detailed product description (feature set)

- **Master Data Management System**
  - Data input & identification
  - 50,000 Shops world wide
  - 5,000 different tape formats
  - thousands of files every day

- **Clients**
  - over 600 customers in Europe, Middle East, Africa
  - served from Nuremberg
  - over 200 Productgroups over 20 countries over customers

- **CONSISTENCY**
  - national / international reports
Stability and Statistical Correctness

Time shifted production of national und international reports implies inconsistencies

operational → *slowly-changing Data Warehouse*

IDAS → DWH → Reporting Base
Final Statement

- there is always a notion of “batch”
  - size may vary from app to app, but |batch| > 1 record

- lookup to ‘approved’ external data sets is OK
  (this has enrichments semantics)

- data flowing through the ETL-process must be carefully handled in terms of local data quality and global data consistency
  -> notion of RELEASE is required!!!
  (release is coupled to data quality assurance, e.g. consistency, completeness)

- I see the requirements of early access to BI data
  - on possible solution: ETL-Query Bus
    - provide a sneak preview of the coming data sets
    - has to be embedded with a notion of quality and consistency
ETL Query Bus

Meta Data Repository

Staging Area

DWH

Data Mart

Data Mart N

incompleteness
data errors
inconsistencies
redundancies

incompleteness

data quality

actuality