Leveraging Information: l'offerta Business Intelligence di IBM

Maurizio Luinetti
maurizio_luinetti@it.ibm.com
Marzo 2000

© IBM 2000

What is Leveraging Information?

To gain competitive advantage customers need to:

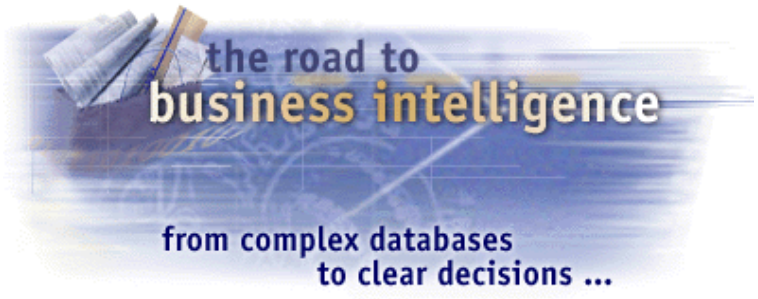
- **Capture data from the transactions and processes of their business**
- **Search & Retrieve structured and unstructured data from all enterprise processes to transform it into meaningful information**
- **Analyze that information to gain deeper understanding of the market, customer needs and internal processes and to turn that understanding into strategic advantage**
- **Provide easy and integrated access, as well as workflow, to that strategic information across the enterprise to share with partners, suppliers and customers**

© IBM 2000

Leveraging Information

Solutions that leverage and manage all forms of information for competitive advantage :

- ***TRANSACTION PROCESSING (OLTP)***
- **BUSINESS INTELLIGENCE**
- ***CONTENT MANAGEMENT***



© IBM 2000

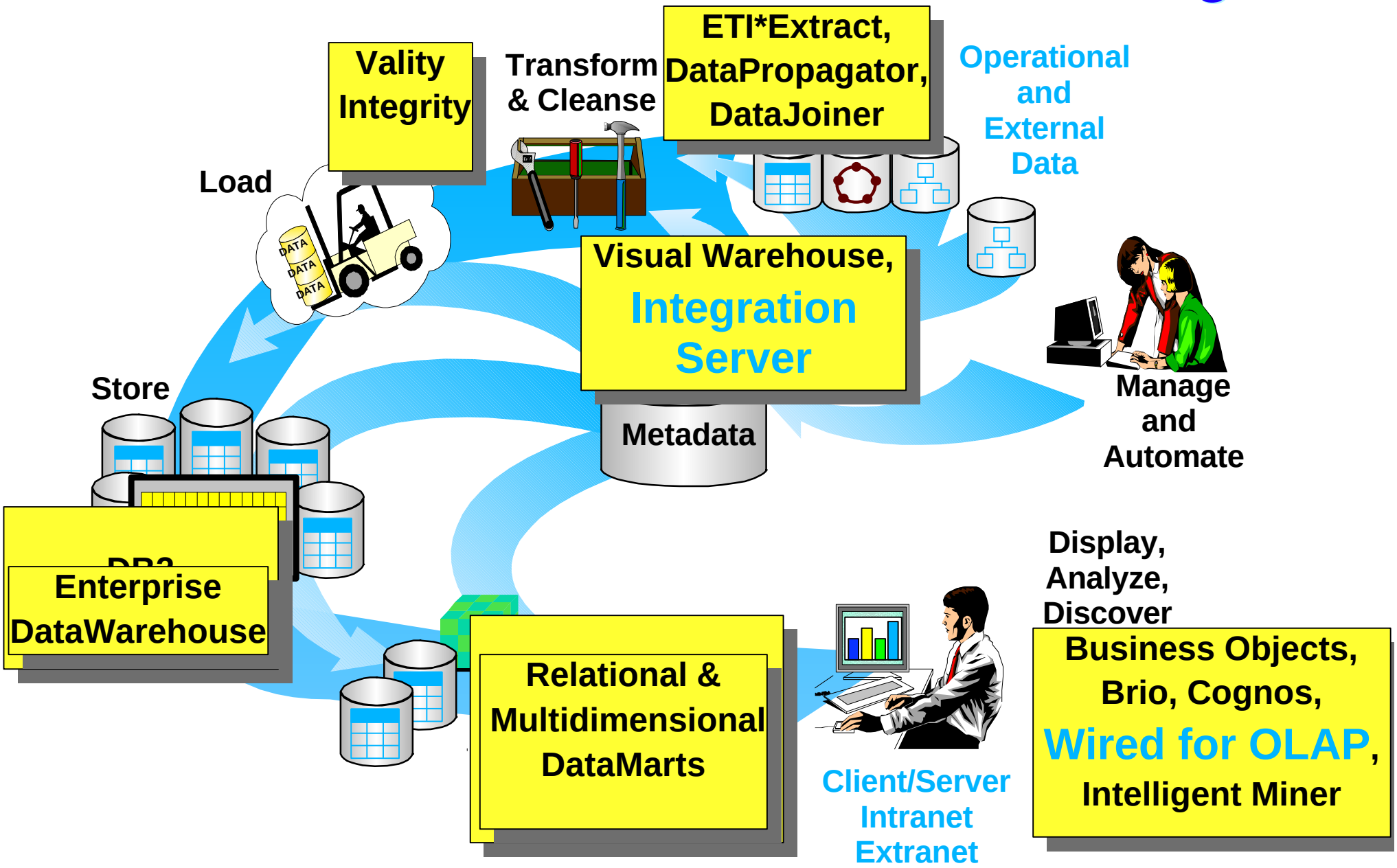
Evolution of Business Information Systems

- **First Generation: Query & Reporting**
  - Host-based system used only by "information providers"
  - Too difficult to use by "information consumers"
- **Second Generation: Data Warehousing**
  - Information is clean and consistent and understandable by business users
  - New user interfaces and more powerful tools
- **Third Generation: Business Intelligence**
  - Focus on access and delivery of business information
  - Information catalog for managing information
  - Support for access to all business information, not only data warehouse
  - Prepackaged application solutions

© IBM 2000

Levinfo.prz

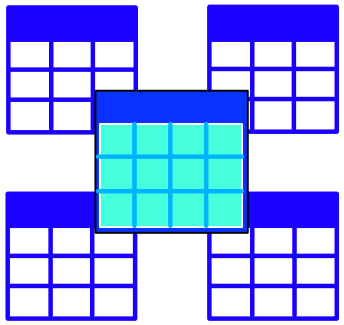
IBM Software for Business Intelligence

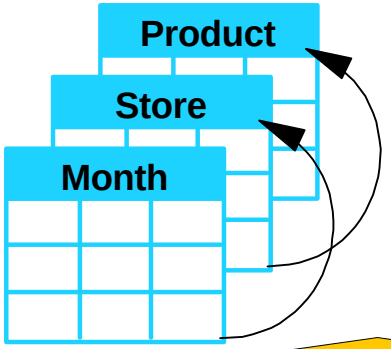


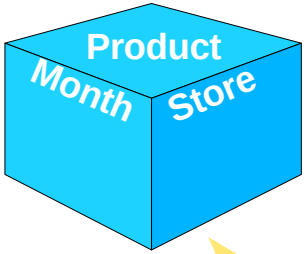
© IBM 2000

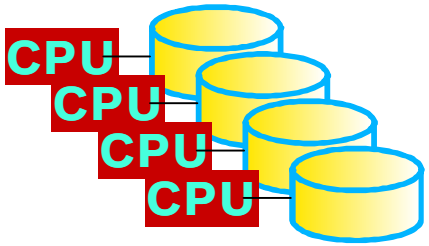
DB2 Business Intelligence Features

- *Star Join*
- *ROLLUP*
- *CUBE*
- *Parallel Query*

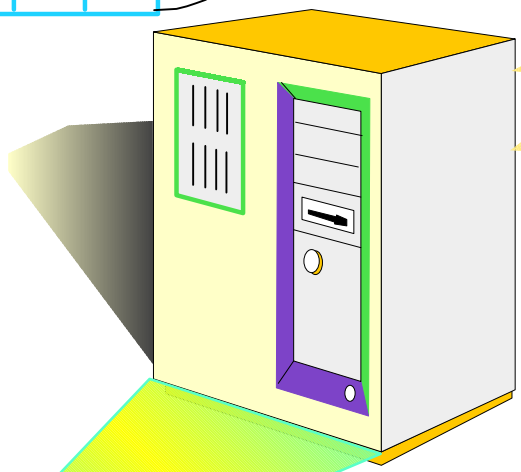








From Laptops
To RISC/6000 SPs



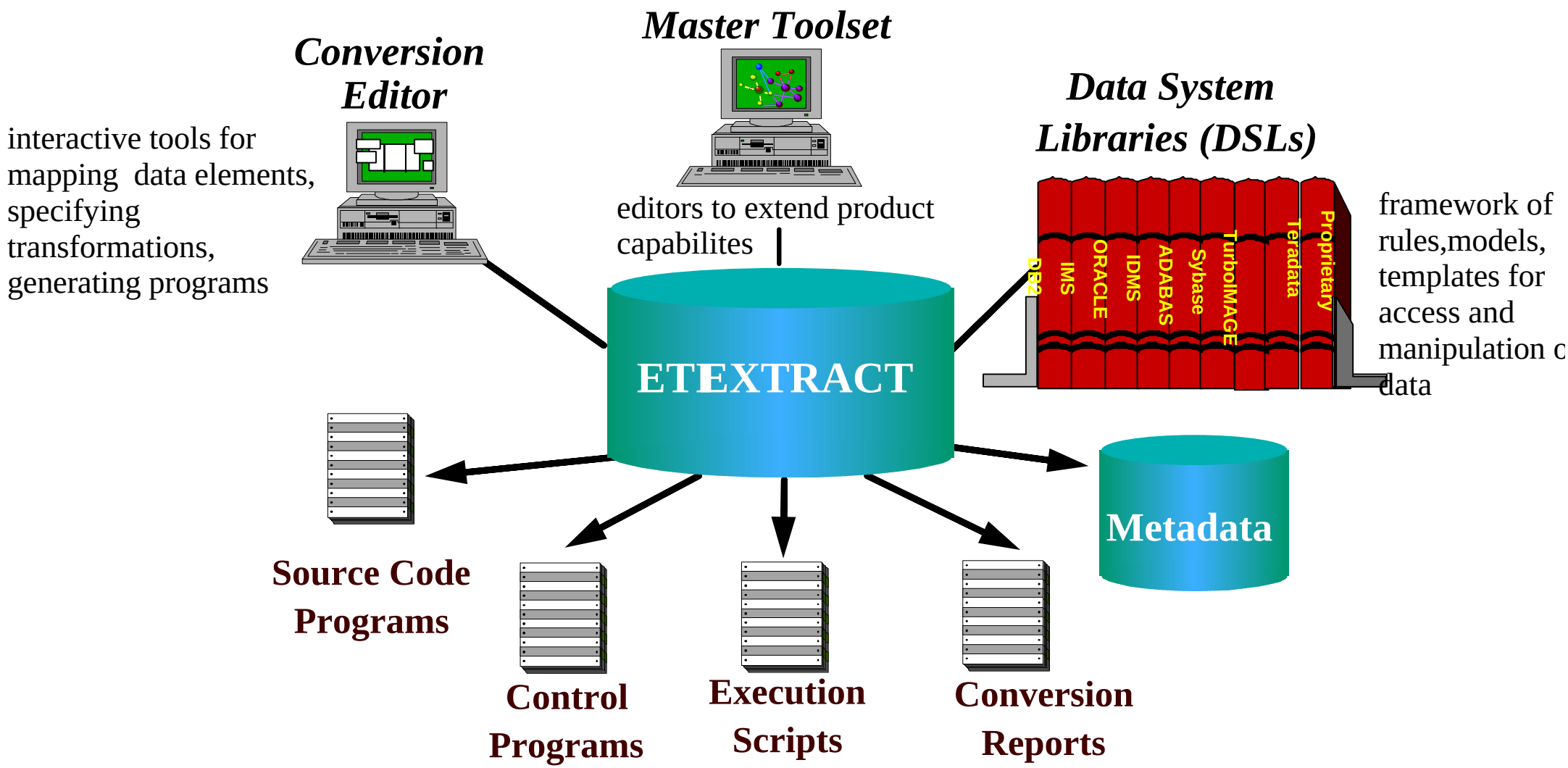
Optimized SQL

GUI-Generated "Ugly" SQL

- *Advanced Cost-Based Optimizer (Starburst)*
- *Query Re-write*
- *Dynamic Bit-map Index ANDing*
- *Automatic Summary Tables*
- *Replicated Tables*
- *Improved indexing*

© IBM 2000

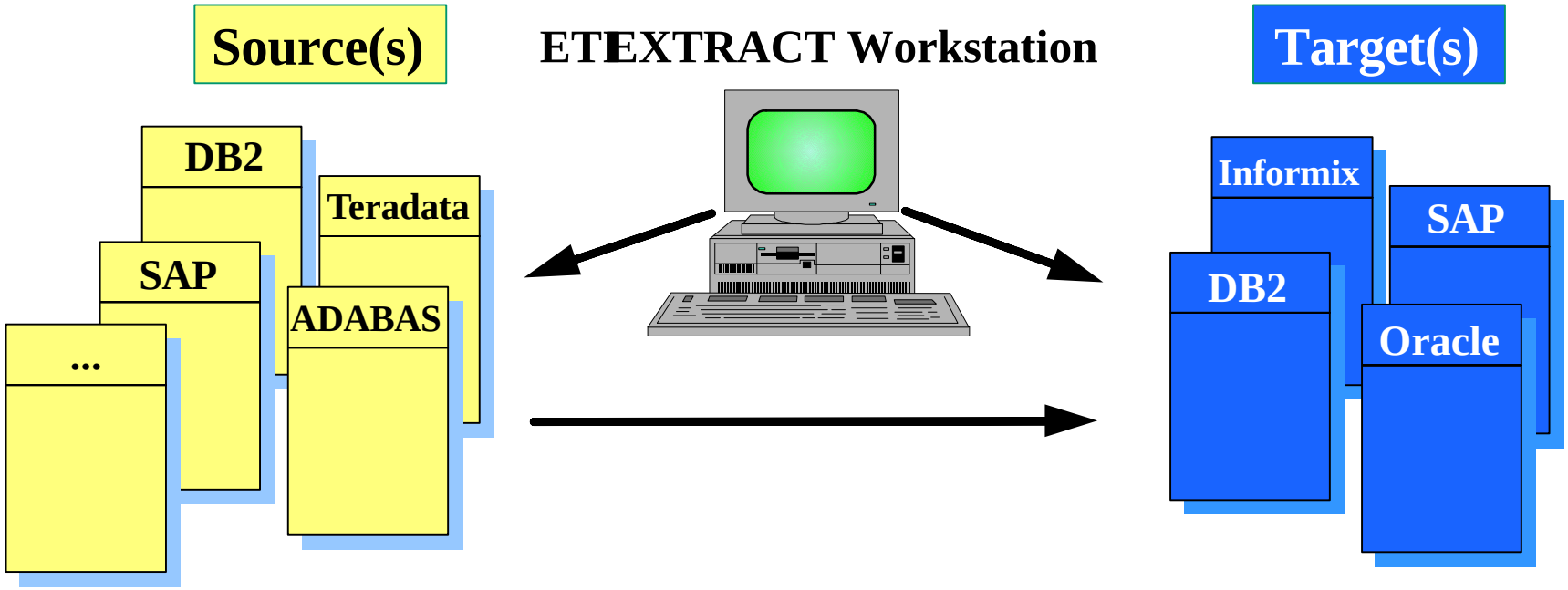
ETEXTRACT Overview



✓ **ETI•EXTRACT automates the writing of programs that retrieve the data needed from any system, transform it, and load it into any other system while capturing the history of that process.**

© IBM 2000

ETEXTRACT Overview



- *Programs generate on ETEXTRACT workstation*
- *Executive transfers programs to source and target platforms*
- *Programs run on source and target systems*
- *Data moves directly between source and target systems*
- *Captures metadata automatically*

© IBM 2000

Vality Integrity

- **Provides Data Re-engineering solutions to organisations requiring high data quality levels**
- **Solve Typical Data Quality problems**
  - Lack of legacy standards
    - Mark Di Lorenzo vs Dilorenzo, Mark
  - Data surprises in individual fields
    - Title in Name field, Fax in Telephone field
  - Legacy myopia
    - Same customer with different spelling
  - Legacy information buried in free form fields
    - Customer information in 3rd line of address
- **Use lexical analysis, pattern processing and statistical matching.**

© IBM 2000

Vality Integrity: Data Re-engineering

External Files

Legacy Applications

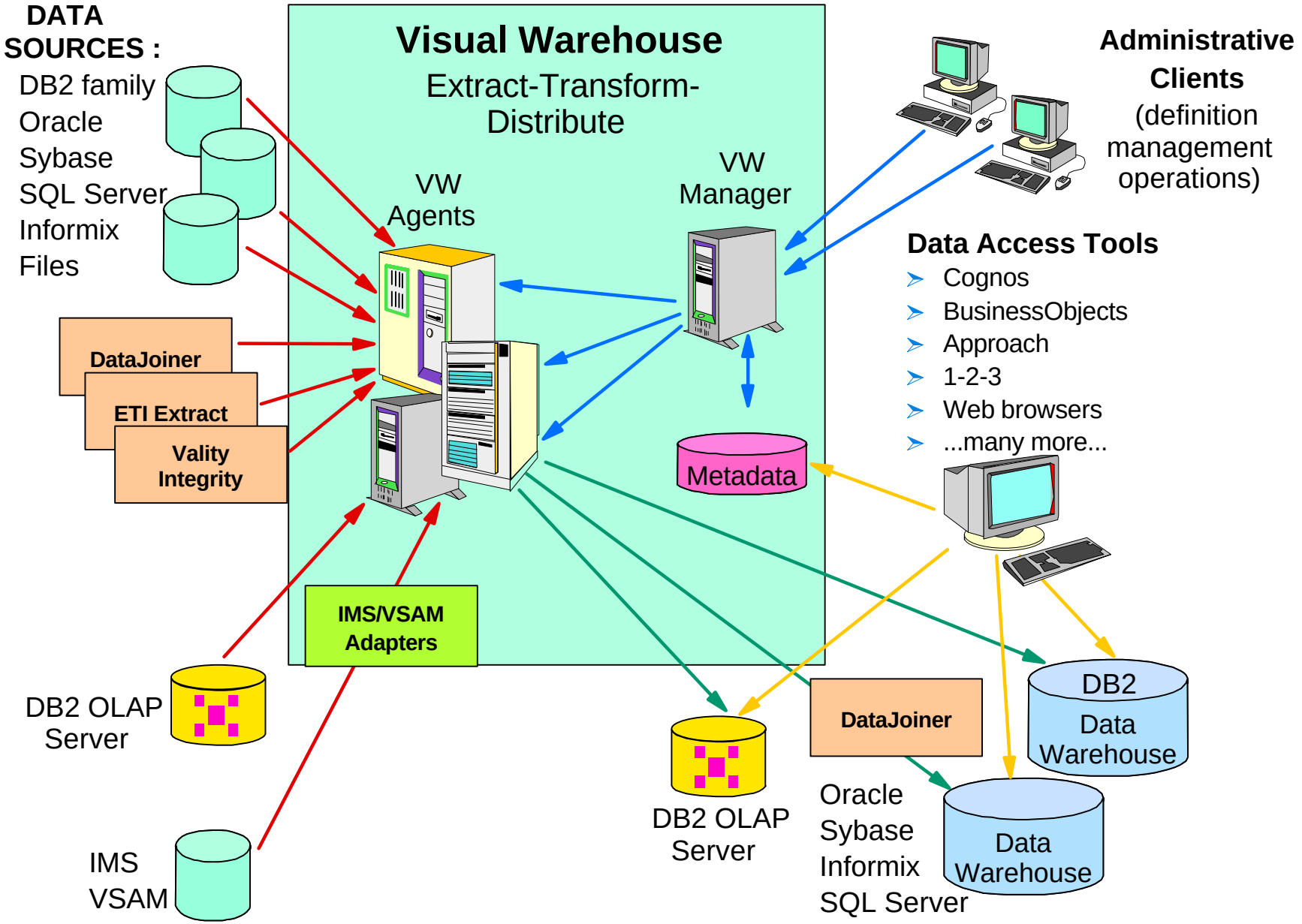
Historical Extracts

1. Data Investigation & Metadata Mining
2. Data Conditioning & Standardization
3. Data Integration
4. Data Survivorship & Formatting

- Data Warehouse
- Data Mining
- Migrations HUON / SAP
- Customer Relationship Management
- Others

© IBM 2000

Visual Warehouse

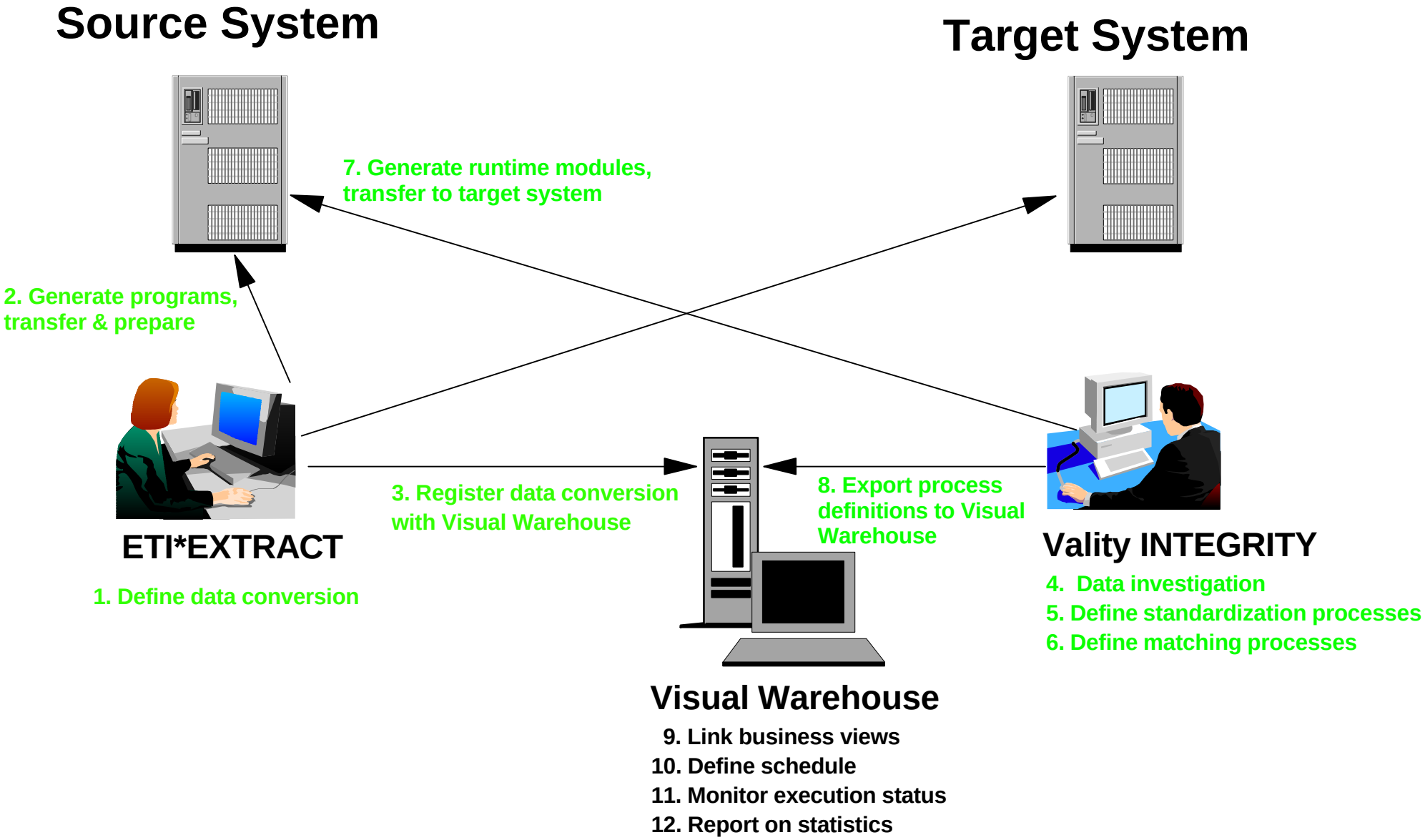


Warehouse Management

- **Define workflow**
  - Scheduled, cascading, program-triggered execution
- **Manage data movement and transformation processes**
  - Visual Warehouse agents (insert or ftp/load), Essbase load processes, ETI generated programs, Vality programs, DataPropagator subscriptions, user-defined programs
- **Synchronize metadata**
- **Monitor execution status**
  - Completed, in-progress, scheduled business views
- **Gather and report warehouse statistics**
  - Automatic statistics collection standard database
  - Summarize data

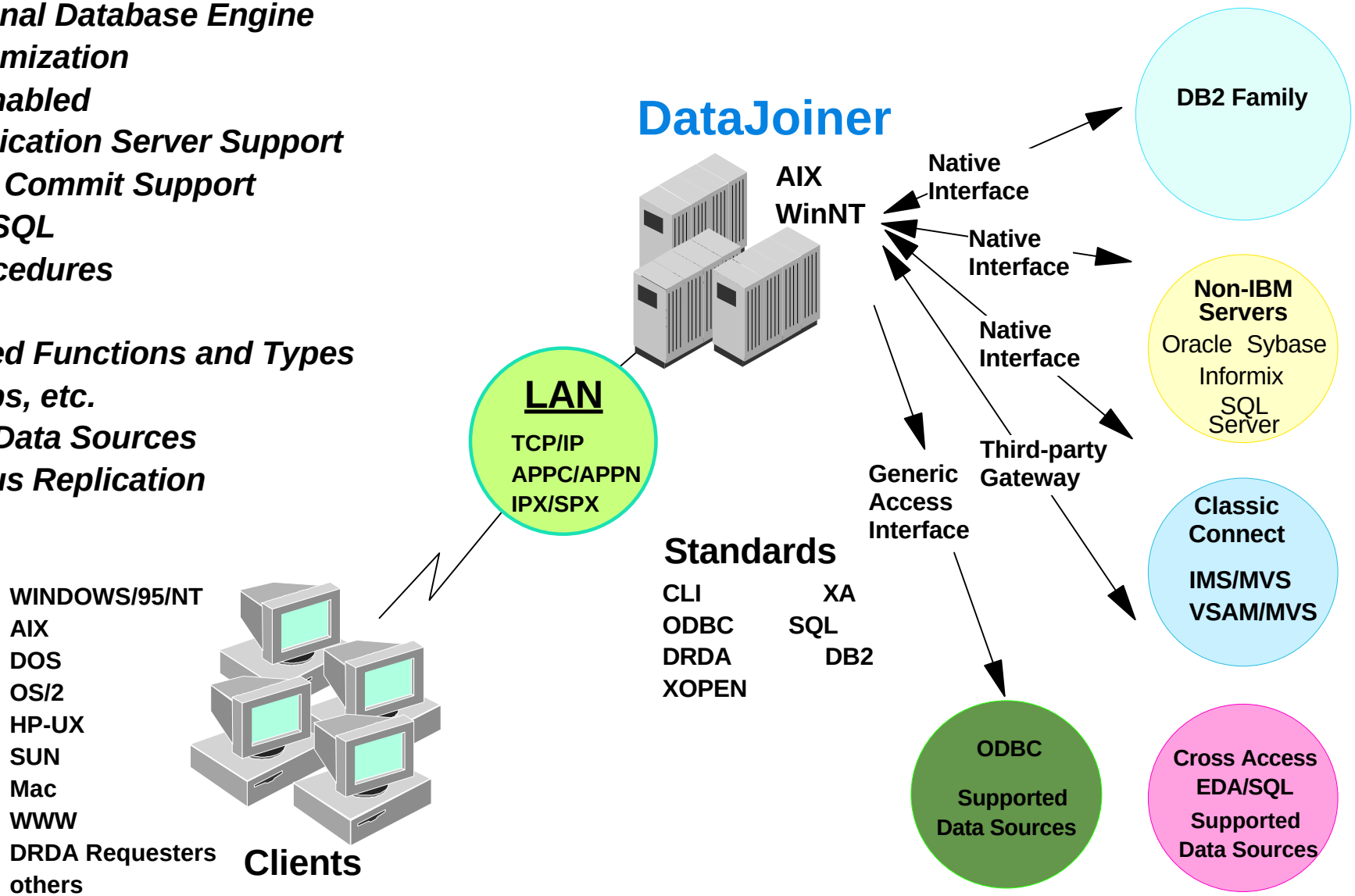
© IBM 2000

Integration with Visual Warehouse

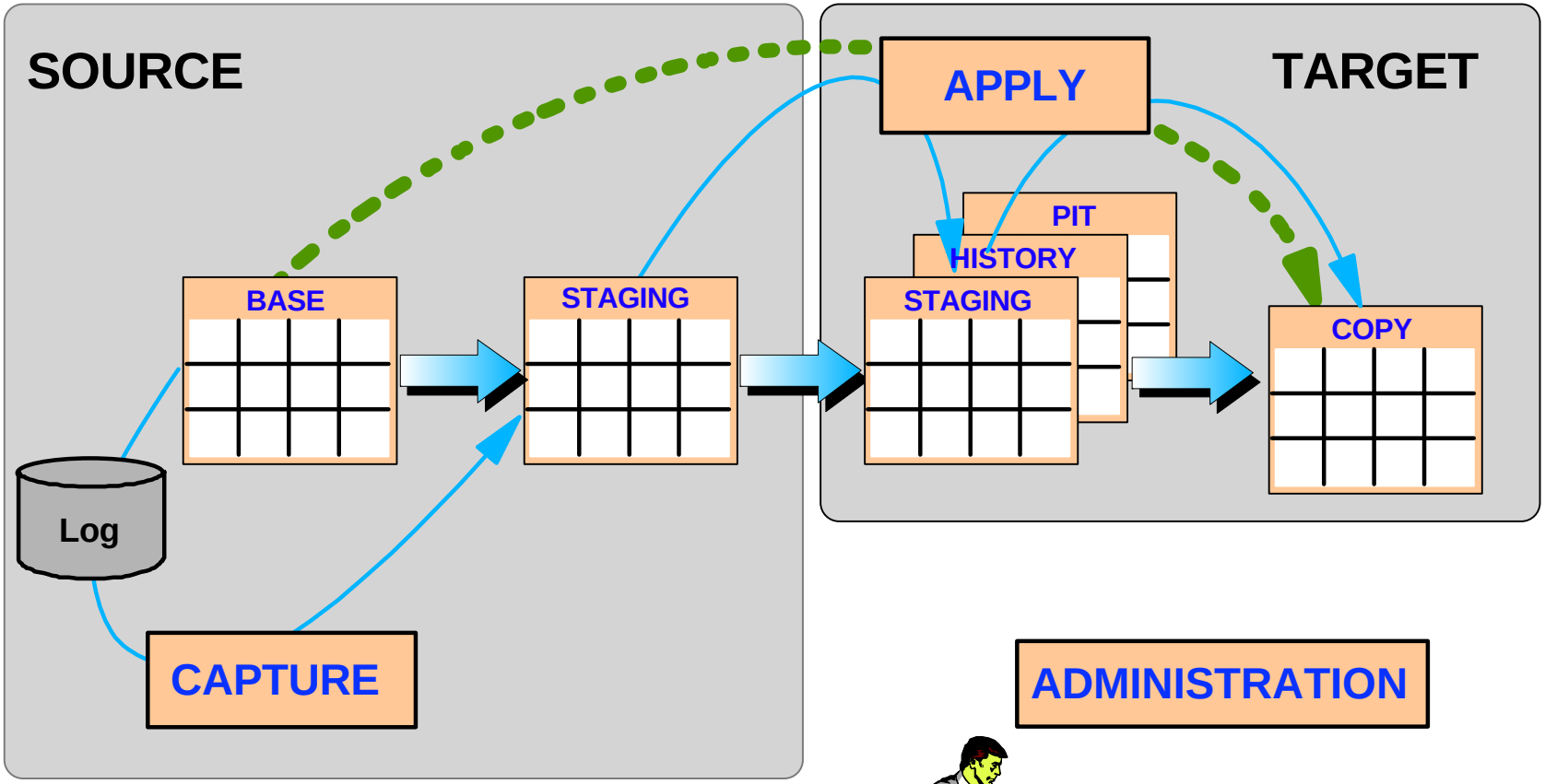


DataJoiner

- *Easy Access/Join Heterogeneous Data*
- *Transparent SQL Read/Write Access to Enterprise Data*
- *Full Relational Database Engine*
- *Global Optimization*
- *Net.Data Enabled*
- *DRDA Application Server Support*
- *Two-Phase Commit Support*
- *Recursive SQL*
- *Stored Procedures*
- *Triggers*
- *User Defined Functions and Types*
- *Blobs, Clobs, etc.*
- *Additional Data Sources*
- *Hetrogenous Replication*



IBM Data Replication - Overview



ADMINISTRATION

DB2 UDB Control Center

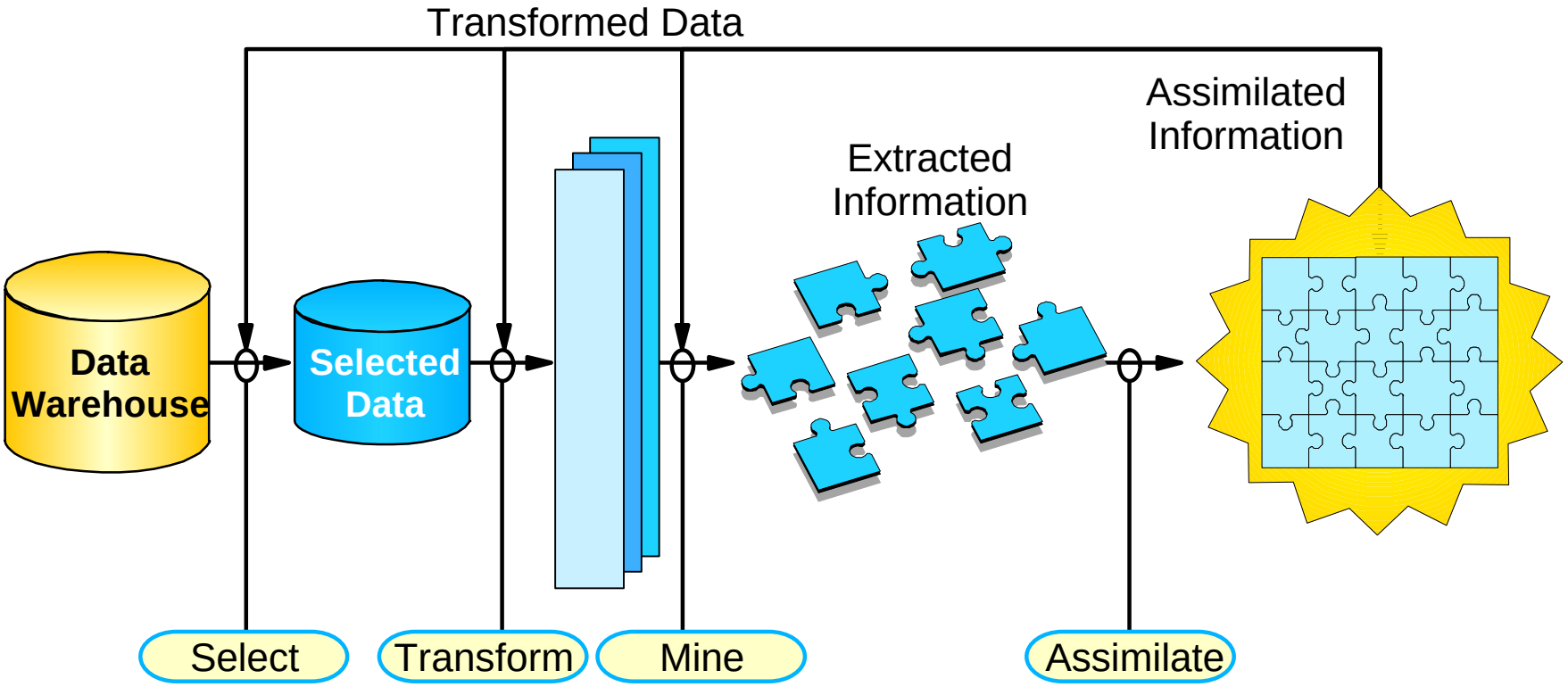
© IBM 2000

IBM Data Replication strenghts

- Homogeneous solution for host and distributed environments
  - investment protection
- DB2 Log based: no impact on performance of operational applications
- replicated data can be manipulated
  - join
  - expressions
  - aggregations
- Condensed or not condensed replication
- Pull or Push configuration
- Bidirectional replication with conflict detection
- Time driven or event driven replication

© IBM 2000

Data Mining is ...



. . . . the process of extracting ***previously unknown***, ***comprehensible***, and ***actionable*** information from large databases and using it to make crucial business decisions.

© IBM 2000

Data Mining Techniques

- **Clustering**
  - Demographic and Neural
    - group a set of individuals based on demographics
- **Predictive Modelling**
  - Classification - Tree Induction and Neural Induction
    - Based on current customers find rules (decision tree), that allow to easily classify a new customer (e.g. credit risk is high, medium or low)
  - Value Prediction - Radial-Basis Functions and Neural Networks
    - Predict likelihood of fraud for a credit card
- **Association Discovery**
  - find items that imply the presence of other items in the same transactions
    - of the shoppers who purchased milk, 55% also purchased some other type of dairy product, and 42% also purchase bread

© IBM 2000

Data Mining Techniques ...

- **Sequential Pattern Discovery**
  - sequence of set of items in the time
    - 42% of new checking account customers who apply for an ATM card in the same transaction, will also apply for our charge account within 90 days
- **Similar Time Sequence Discovery**
  - sequences similar to a given sequence in a database of time-series data
    - given "NorthSouth Airlines" daily closing price for 1995, find all other stocks with similar behavior

© IBM 2000

Data Mining : IM for relationship marketing

Customer Attrition

Builds a predictive model on the probability a current customer will leave.
Profitability analysis
Gain charts for profitability cutoff

New Customer Acquisition

Builds a predictive model of positive respondents to a marketing campaign.
Profitability is included in the model

Customer Winback

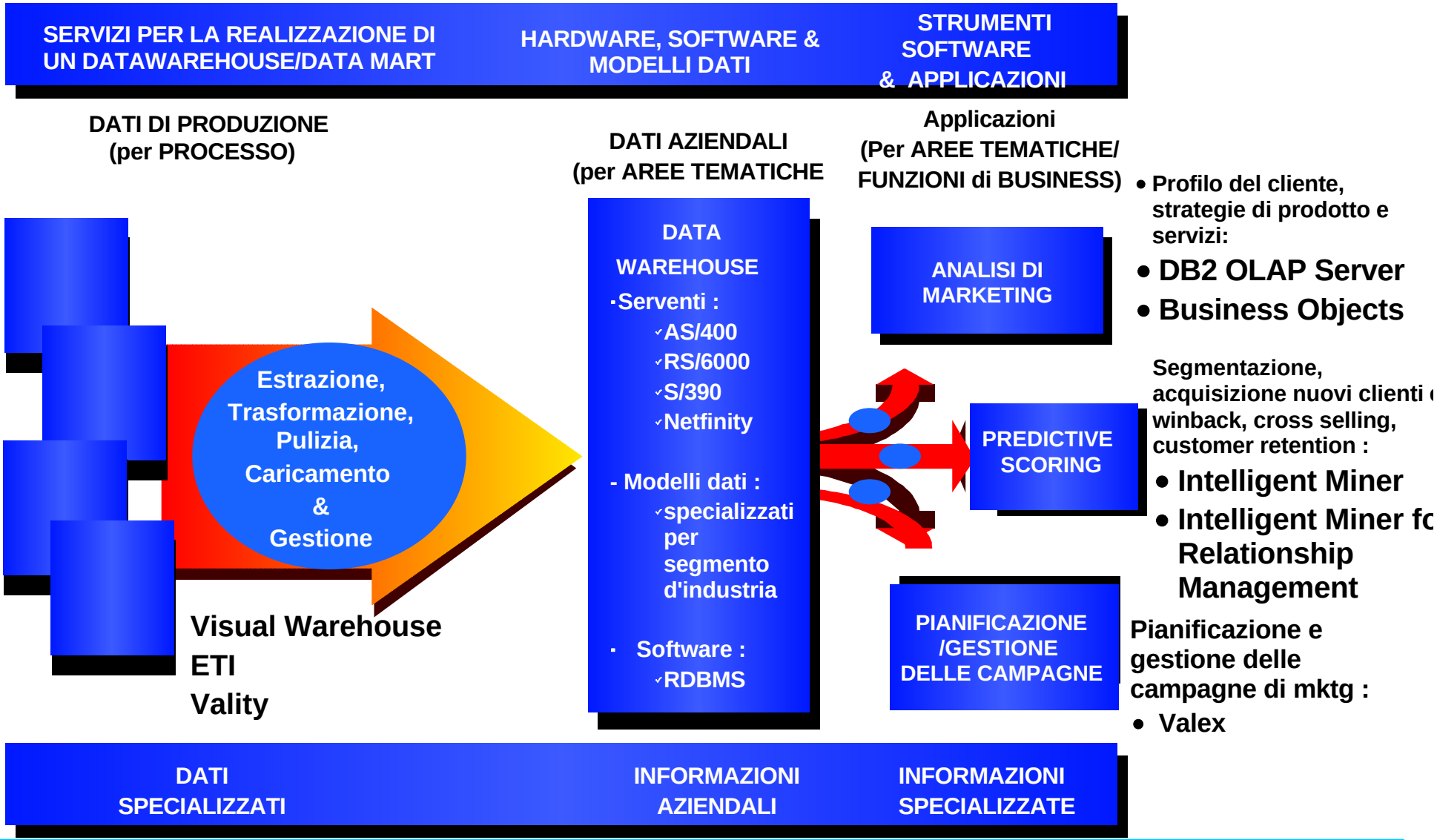
Uses above to build a model of former customers to winback. Uses data to avoid winning back unprofitable customers, and target favorable customers.

Cross Selling

Uses association rules for each segment to identify opportunities to sell multiple lines in retail, banks and insurance.

© IBM 2000

Relationship Marketing Architecture : Decision Edge for Relationship Marketing



© IBM 2000