Beyond the Data Warehouse

Data Warehouse: A Mandatory Initiative For Survival in the 1990s

Aaron Zornes
Senior Vice President
Application Development Strategies
META Group

Notes
Beyond the Data Warehouse

Data Warehouse (DW) Definition

A subject-oriented information store, designed specifically for decision support

Types of information store
- Relational database (RDBMS)
- Multidimensional database (MDBMS)

Notes
- DW does not automatically imply relational; DW can be distributed across multiple file types.
- IT organizations are under tremendous pressure to provide better quality, decision making information in forms easy to access and manipulate. Business users are reacting to their own "personally mission critical" need for better information due to: (1) rapidly changing, increasingly competitive markets, (2) increasingly volatile consumer and market behavior, and (3) rapidly shortening product life cycles.
- A 1994 META Group survey of IT managers at Fortune 2000 enterprises found that over 90% were planning to implement data warehouses in 1994-96. Clearly in these times of increasingly restrictive IT budgets, either lines of business are championing these next-generation decision support systems (increasingly referred to as on-line analytical processing or OLAP), --or-- IT organizations are investing in projects believed to have major, demonstrable near term payback. META Group believes it is both.
Notes

Data warehouse (DW) = a subject-oriented information store (relational DBMS or multidimensional DBMS) designed specifically for decision support. DW objective = provide an integrated framework for delivering corporate info to knowledge workers. Data warehouse is a blend of technologies – relational and multi-dimensional databases, client/server, graphical user interfaces, powerful metadata modeling, and more – to enable the integration of multiple operational databases into a single database designed specifically for analytical processing (e.g., decision support). The resultant "subject-oriented database" is designed with end user access in mind.

DW role = (1) reconcile sources for all EIS/DSS data, (2) minimize/eliminate extract processing, (3) integrate historical data & provide foundation for EIS/DSS applications.

Key differences between data jailhouses (operational databases) & data warehouse databases:
- Subject orientation – operational systems are application-segmented (i.e. for banks: auto loan, demand deposit accounting & mortgages). Subject areas for banks would be "customer" & each "financial product."
- Level of integration – data warehouses resolve years of application inconsistency in encoding/decoding, data name rationalization, etc.
- Update volatility – record-at-a-time updates in operational databases vs. bulk loads in data warehouse
- Time variance norms include: 30-90 days of transactions for operational system, 1-5 years for data warehouses.
Beyond the Data Warehouse

On-Line Analytical Processing (OLAP)

- Variation of decision support to manage complex analysis of "dimensional" data
  - Sales (product, size, region, time)
- Requires a back-end data store which may be
  - Physical MDBMS
  - RDBMS with logical dimensional metadata
  - Specialized file structure

Notes

- OLAP is a subset of decision support
- OLAP back ends are data warehouses
Notes

▲ Today
- "Reactive" query governors
- Intraenterprise (enterprise-owned) data
- Security via database object (view, table)
- Polling style execution
- System failure = execution cutoff

▲ 1995 & Beyond
- "Prequery" costing
- Extended-enterprise boundaries (on-line services, etc.)
- Security at business object level
- Active, agent-style execution
- Middleware-enabled, guaranteed delivery
Beyond the Data Warehouse

Data Warehouse “Drivers”/ Business Issues

▲ Organizational
  ● Enterprise downsizing/rightsizing

▲ Operational
  ● Vertically challenged, application-specific databases
  ● Differing versions of the “truth”

▲ Political
  ● IT hegemony vs. end-user ad-hocracy

Notes

▲ Organizational drivers
  ✦ Distributed decision support mandated by enterprise downsizing/rightsizing
  ✦ More DSS/EIS capabilities required – often unpredictable, heuristic, & subject-oriented
  ✦ Decision-making flexibility enabled via multitiered data structuring – from detailed, transactional level to high-level summary

▲ Operational drivers
  ✦ Need to fix poor data availability, consistency, & integration of operational databases
  ✦ Inadvisability of manual extracts against production OLTP databases

▲ Political drivers
  ✦ Need to reinvigorate IT’s relationship with end users

▲ Why is “Data Warehouse” one of the hottest IT trends? Enterprises must become more competitive in order to survive. Increasingly few markets or industries enjoy monopolistic advantages. Airlines, utilities, and health care organizations are but a few of the businesses which must increasingly compete both in terms of cost and service – even government agencies find themselves competing against the specter of external outsourcing. To compete in the 1990s, enterprises must get closer to their primary markets. Unfortunately, much of the data maintained by enterprises about their customers are locked up in “data jailhouses” – databases which have evolved as discrete independent subsystems. These “vertically-challenged” databases are unable to provide the enterprise a consolidated view of who the customer is, or even what services and products are correlated across the customer base.
Beyond the Data Warehouse

Justifying The Business Value

- More cost-effective decision making
- Better business intelligence
- Enhanced customer service

Notes

▲ More Cost-Effective Decision Making – Elimination of staff and computer resources required to support ad hoc inquiry and reporting against operational/production databases offers significant savings. "Profitability analysis" is also a key benefit of being able to correlate combinations of products and services with marketing initiatives and external on-line business demographic databases.

▲ Better Business Intelligence – Increased quality and flexibility of market analysis arises from multi-tiered data structures which may range from detailed, transactional level to high-level summary. This bypasses vertically-challenged, application-specific databases which foster end user and management distrust due to differing versions of "the truth". The ability to discern "how many" of "which products" are actually sold into a given household address ("householding") is key to identifying cross marketing opportunities as well as providing critical insight into the business's best customers.

▲ Enhanced Customer Service – The total customer relationship can be established if all information concerning a customer can be correlated via a single data warehouse.
Justifying The Business Value

- Enhanced asset/liability management
- Business process re-engineering
- Alignment with corporate downsizing objectives

Notes

- Enhanced Asset/Liability Management – Purchasing agents and other financial managers could benefit greatly if they were to ascertain the "big picture" of enterprise-wide purchasing and inventory patterns. In doing so, the financial managers often discover cost savings hidden in redundant inventory, as well as identify previously unknown volume discount opportunities.

- Business Process Re-Engineering – Providing business users unlimited analytical analysis of their business information often provides insights into the work processes themselves and will yield breakthrough ideas for the re-engineering of those business processes.

- Alignment With Corporate Downsizing Objectives – Distributed decision support is increasingly mandated by enterprise downsizing/rightsizing a.k.a. "corporate bulimia" as organizational restructuring further distributes decision making responsibility.
Beyond the Data Warehouse

Key Trend – Rate of Investment

Figure 1: Overall DW Investment

Notes

△ The average company responding has already spent approximately US$3M (project to date) on hardware and software (hardware investment averages nearly US$2M). In aggregate, over US$300M has been spent on DW projects.
Beyond the Data Warehouse

Key Trend – Very Large Data Bases

Figure 2: Warehouse Database Size

Notes

▲ Among our sample, 19% of respondents are already at the 50GB level, 59% expect to be there by 2Q96.
▲ DW size increases not necessarily as longer periodicity is designed in; rather, DW is an incremental process whereby IT adds additional subjects and department data.
▲ The objective is to provide quick incremental results.
Beyond the Data Warehouse

Key Trend – Explosive Growth in Users

Figure 3: Estimated Number of DW Users

Notes

▲ Although very few DW projects today support more than 50 concurrent users, 83% claimed they would be supporting 50+ users by mid-1996.

▲ Number of DW users also reflects incremental process whereby IT adds additional subjects and department data.
Beyond the Data Warehouse

Key Trend – End-User DSS Market Leaders

- MS Access – 11%
- Focus – 8%
- SAS – 7%
- QMF – 6%
- PowerViewer – 6%
- Forest & Trees – 6%
- Paradox – 5%
- Pilot – 4%
- Clear Access – 4%
- Business Objects – 4%

Notes

▲ With respect to RDBMS-oriented end-user tools, ubiquitous spreadsheets and Microsoft Access dominate in terms of tools currently used, but moving forward, are being evaluated less often by our survey respondents than modern queryEIS tools (e.g., Business Objects, Cognos' Impromptu/PowerPlay, Platinum Technology's Forest & Trees, and MicoStrategy's DSS Agent).

▲ SAS also provides data management and data organization (extracting and loading into DWs, storage management for DWs) and is therefore also a more robust DW solution than standalone DSS tools.
Beyond the Data Warehouse

Data Warehouse "Enablers"/ Technology Issues

△ Metadata
- DSS-oriented metadata modeling capabilities – including scrubbing/cleaning/enrichment
- Business user-oriented repository – a.k.a. "Information Catalog"

△ Data delivery
- Cost-effective midrange Unix & LAN RDBMS servers
- Data replication/copy management facilities
- Standardized database access middleware
- Highly compelling, C/S-enabled DSS/EIS

Notes

"Key" IT considerations:
"Information Architecture" plan
Because no single software vendor provides integrated, off-the-shelf solution, IT turns either to: (1) Big 6 consultancies, (2) systems vendors professional services, or (3) IT internally integrates components.

© 1995 META Group, Inc., Stamford, CT (203) 973-6700
Beyond the Data Warehouse

Getting From Here To There

▲ Assess key business dimensions ("subject areas")
  • Identify prime candidate "knowledge workers"
  • Hold JAD/RAD sessions
▲ Estimate and locate ("data archeology")
  • Perform data sizing projections
  • Identify lightly vs. highly summarized aggregation
  • Identify periodicity requirements
  • Begin archaeology search for data sources
▲ Pilot and production introduction
  • Extract, cleanse and load data
  • Demo and solicit feedback
  • Meet and exceed current needs
  • Plan for the future

Notes

▲ Detailed JAD sessions
▲ Identify sources, targets, mapping
▲ Determine data refresh frequency
▲ Highlight degrees of "dimensionality"

© 1995 META Group, Inc., Stamford, CT (203) 973-6700
META Trends – Strategic Planning Assumptions

During 1995/96, data warehouse (DW) architectures will enable component-level integration of OLAP access with corporate OLTP applications and data.

Through 1996/97, key challenges for large scale DWs include lagging support for metadata synchronization, information catalogs, and DW-smart database design tools and methodologies.

During 1995/96, DSS/EIS workbenches will mandate support for both remote SQL query tools and local multi-dimensional data marts.

These next-generation DSS/EIS will be enabled by external and "soft" information sources, coupled with metadata-driven intelligent agents (1996/97).

Notes
Beyond the Data Warehouse

Summary – Bottom Line

▲ Data warehouse is an essential IT initiative
  • Leverages existing technologies
  • Defines new business opportunities
  • Changes the basis of competition
▲ IT must
  • Provide architectural guidance via an Information Architecture
  • Start with small pilot with focus on decision making “pain points”
▲ During 1995/96, changed data capture, workflow and information catalog capability = key differentiators

Notes
▲ During 1995/96, META Group analysts believe the majority of corporate IT organizations will re-engineer DSS/EIS systems via formal OLAP and data warehouse architectures and initiatives. Whether it be by providing marketing insights, opportunities for cost savings or volume discounts, etc. – clearly, data warehouses provide an increasingly critical component of the information systems being delivered to support the increased tempo of business competition.
Beyond the Data Warehouse

Q & A

Notes