

BUILDING A BANKING CUSTOMER RELATIONSHIP DATA WAREHOUSE: A CASE STUDY

Noor Quadri, Oracle Corporation

INTRODUCTION

This case study centers on a large banking organization destined to develop a customer relationship data warehouse in order to meet competitive demands and improve its customer service and profitability. Key business activities, scoping process leading to development of the Data Warehouse will be discussed along with reference to technical architecture, but, not the data base layout and related metrics owing to paper space limitations. Due to competitive nature of financial business, name of the Bank in question will not be disclosed.

BANKS POSITION

Bank's commercial and investment activities dominate a vigorous and highly competitive marketplace. The asset base currently exceeds US\$ 2.9 billion and its loan portfolio accounts for US\$ 1.2 billion.

The bank operates over 80 branches locally and internationally. It provides a comprehensive range of products and services to its private and corporate customers. These include current, savings and fixed deposit accounts, loan and overdraft facilities, foreign exchange currency services and documentary bills and credits. The Bank, through its subsidiary companies, provides financial services in the areas of off-shore banking, life assurance, fund management, house finance services, flexible manufacturing and export loan facilities, as well as long-term finance for industrial projects.

THE DATA WAREHOUSE STRATEGY

The central strategy driving the Bank has been that of transforming the core banking business into more diversified range of activities. The bank identified its competitive advantages as being its strong distribution network, its acceptance as an established institution, its financial strength and the ability of its personnel to service the local and expanding international market. New channels of distribution complementing the branch network were, therefore, required. Product diversification plan was complemented by a strategy of facilitating the manner in which customers access banking services through delivery of Data Warehouse engineered architecture.

The Bank therefore decided to develop the technology framework in order to deliver the relationship marketing strategy that is needed to sustain its business objectives aimed at offering services directly to customers. A critical success factor underpinning the strategy was the synergy between the Bank's marketing and information technology resources. Relationship banking through database marketing was identified as the solution to merge the Bank's marketing and information technology capabilities. Strategic value of the data warehouse was understood as key to successful deployment of customer service components and product profitability.

The bank embarked on the Data Warehouse strategy based upon series of one-to-one meetings with senior management from various departments. Direct marketing met the fundamental business needs of the Bank's short-term strategy to convert branches into selling points. Relationship banking was pointed out as the corner stone of the bank's I.T culture in achieving the depth of knowledge that this strategic thinking requires over the medium and long-term. This was achieved through exploring the use of data warehousing technology, which raised the ability of all employees and decision workers to serve the Bank's customers. The greater the amount of available data and the number of employees with access, the greater the strategic leverage the Bank will receive from its Data Warehouse solution through effective information management. Further more, the bank wide dissemination of information was required to be propagated through the intranet infrastructure in order to be effectively introduced in time for the data warehouse development. Such data classification was missing in the Bank's information systems. This strategy served as a platform for both the Data Warehouse as well as to support the migration towards the new retail system and wholesale on-line transaction processing banking system that Bank envisioned for the future.

The basic component of bank's marketing strategy was to support profitable customers. When customer files are matched with customer calls, an ordinary telephone conversation could turn into a productive sales session, while the customer appreciates the Bank employee's knowledge of the account, who can concentrate on the task at hand. This

database marketing initiative relating to customer knowledge, buying habits, use of banking products was singled out as the first building block towards the preparation of a prototype plan/proof of concept for the Bank's data warehousing initiative. Therefore, the key performance indicators requiring immediate attention and to draw executive management interest in such a project was deemed necessary by addressing the following critical areas:

- Matching customer needs with the appropriate banking services or products
- Identify cross-selling opportunities for the Bank and its subsidiaries
- Build a relationship banking information architecture

By giving its decision makers robust access to information about customers, markets, suppliers, financial results and products/services, the Bank will empower its people to strategically learn from the past, adapt in the present and position for the future and hence Data warehousing was seen as an opportunity to set out a framework showing how the marketing units will work closer together to co-ordinate the Bank's database marketing initiatives. The Data Warehouse framework will enable both Business Development and Direct Marketing officials to enhance their understanding of data warehousing concepts, and set in motion a tactical plan to support the Bank's efforts in merging its marketing and IT capabilities.

SCOPING STUDY AND RESULTS

One of the fundamental milestones of any Data Warehousing engagement is the collection of business requirements and Organization's commitment to the project with appointment of a project sponsor. Therefore, as a primary step, scoping study performed resulted in definition of the Data Warehouse blue print and the road map. This study provided advice to an appropriate program of work to achieve the defined business objectives and to offer Oracle's support in achieving this initiative. In line with our recommendation the approach deployed to understand the direction and strategic thinking of the Bank was to conduct a detailed and very comprehensive user workshop in understanding bank's real needs and gaining complete understanding and commitment to deliverables by bank's executives and Data Warehouse sponsor. Once this was achieved, the strategic logic of the data warehousing proof of concept became clearer and the path to an optimum implementation emerged.. From a business perspective, the Bank expected Data warehouse deliverable to provide a common view of Customer and its related functions, products/services used etc regardless of the underlying architecture. A customer is a customer is a customer, regardless of who is asking the question. Such a shared environment will provided the Bank with more accurate and complete information for better decision making. This single view of all the information stored about each individual customer will allow the Bank to provide timely, decision making information concerning customer and financial services trends.

The Scoping Study was conducted through series of interviews and/or workshops aimed at identifying bank's underlying business objectives. The study also reviewed existing work in the areas of Business Process Improvement, IS Effectiveness and Change Management, identifying and prioritizing areas of risk. Distinct deliverables are discussed below:

Deliverable	Description
Data Analysis and modeling	Production of logical data model and specification of key processes and specifications on Relationship banking deliverables
High level technical architecture.	Technical architecture that fits Bank's Data Warehouse requirements.
Project plan	Project plan and roles/responsibilities
Risks and issues.	Identified risks or as discovered during the scoping study with suggested mitigation strategies.

FORMAT OF THE SCOPING STUDY

Discovery process provided fundamental input into business requirements, business issues etc. End user s,I.T, departmental managers, spent over two weeks in a rigid planning session environment with leadership provided from the business sponsor. Following pertinent details recaps this phase.

- Agree on terms of reference
- Agree on Strategic direction and requirements
- Analyze business user requirements and Source Systems
- Review infrastructure
- Review operational systems
- Review existing reporting and Decision Support Systems
- Develop high level project plan and review risks
- Discuss oracle warehouse methodology
- Study the Current infra structure, data sources, technical architecture
- Understanding of Corporate Vision
- Critical Success Factors, Business Drivers
- Industry Information - compliance, mandates, government initiatives
- Business Goals and Strategy
- Organizational Responsibilities
- Products, Channels , Customers
- Business Issues & Drivers
- Cost Benefit Analysis or Return Business Value Analysis
- Major Areas of Benefit
- Conceptual Data Model
- Analytical Requirements
- Current Transactional Systems Evaluation of source transaction systems in relation to decision support goals:
- Data requirements, data availability, time windows, Source Data mapping needs
- Infrastructure Review, Forecast & Basic Gap Analysis: Client Systems, Server systems & networks

SAMPLE OF QUESTIONS ASKED DURING THE SCOPING STUDY

Hopes,Fears,Expectations, How many data elements ?

Missing data from Source and LDM, Is gap analysis report available?. If not, are there resources available who can help identify

All business rules clearly defined to scrub, clean and load, Met data requirements,Data transfer, refresh window ?

Would Bank allow changes to existing systems if needed?

Are any of the application data time stamped ?

Can Bank confirm business subject area and who are the users, Power, End User, Management and any service level requirements?

Is Bank open to data storage/model 3NF,Star schema?

Is automated change capture mandatory ?

Is name/address transformation required ?

List type of other transformations required ?

Delete, Insert, Change policy in source systems

Does Bank have an identified LDM and business process(s) identified?

Are domain codes fixed in source system. Would Bank want to see similar codes in DWH ?

Are segmentation codes available?

Above list of activities is just the tip of the iceberg and detailed interview sessions lead to greater understanding of Banks plan to use Information from the Data Warehouse and to develop a detailed LDM and architecture.

The Data Warehouse delivery service levels required warehousing project to offer the opportunity for decision makers of the Bank to have access to primary and secondary information of customers and its relationship in context of marketing and profitability needs. The user interface was a GUI based.

Different corporate decision makers required on the fly access to set of information that needed to plug in data from several OLTP banking systems. This requirement necessitated a rich data access layer with intricate dimensional data modeling exercise. To be specific, the following category of users required access to data as follows:

- The operational user, with a requirement for frequent reports that can easily be pre-defined, such as branch administration staff, marketing staff and head office departmental staff.
- The power user is the sophisticated user, such as marketing and brand coordinators and financial officers, who have a range of PC skills and are capable of generating ad hoc queries.
- The Chairman, CEO, GMs and other members of the executive branch, were supported by professional query personnel, generating SQL to serve the purpose, with an Executive Information Dashboard interface that allows the executive user to easily drill-down into the detail.

During the scoping workshops, several proactive business processes were studied leading to some of the following questions from business users

- How can we identify, develop and exploit profitable customer relationships?
- How can we create a product and service offering tailored for those customers' dynamic needs?
- What is the right distribution system for developing long term competitive advantage?
- How can we manage risk while serving customers and sustaining superior returns?
- How can we align and coordinate the actions of our employees to maximize performance?

KEY PERFORMANCE INDICATORS

The following areas of critical business impact were discovered out of a vast requirement base:

- Relationship banking
- Cross segment marketing
- Risk and credit analysis
- Industry sector analysis
- Customer Profiles
- Branch Performance
- Product Portfolio and
- Profitability

Some of the following(Sampling) KPI's(Key Performance Indicators) provided in depth information flow and business requirements and assisted in development of a logical data model(LDM)

INITIATED BY	KEY PERFORMANCE INDICATOR
CEO	Single statement Banking
CEO	Measure net Interest Income
Marketing	Credit Scoring
Marketing	The transactional penetration in each product by each customer is a geographical region.
Marketing	The product basket for each customer in a geographic region..
CEO	The most prominent life cycle segments being serviced by each branch
Marketing	Tracking on specific Customer groups
Treasury	The penetration in the transaction history of each customer
Marketing	The percentage of the customer's wallet serviced by the Bank.

SAMPLE OF TRANSFORMATION RULES

The following table shows one of the outputs of the scoping study with an agreement on transformation rules:

Validation Requirement	Process
Type Checking	Data should conform to business rule definition Data should fall within acceptable values
Domain Checks.	Reasonableness checks. Summary totals
Integrity Checks	Ensure table and data relationships are consistent
Business Rules	Ensure data is correct in context of business context Example: If NSF then add 1 to credit score.

DATA LOAD RULES

The following load rules were agreed during the workshops:

- Error logging during load is mandatory
- Hash controls for data loads
- User notification for any refreshes to data
- When errors occur, manual checks to be deployed. Fix errors and continue
- Data to be time stamped and purge frequency determined accordingly
- Complex transformation were built into load process. Example ,computation of average daily balance, Full text description f addresses
- Data scrubbing was part of the pre load process.

SOME DESIGN DETAILS

De normalization process was used where appropriate in the architecture. This provided simplification of queries with the eliminating of joins and repetitious aggregations. This technique was used for code look up and other process details. Code description was used instead of storing some of the most common codes. This method although took

more space, but, resulted in much improved access and response times. Load process was customized to check certain codes and replace by its proper descriptions.

The warehouse design deployed controlled duplication and aggregation of data to reduce the number of tables accessed, eliminate aggregations and speed up user queries.

Changing dimensions design was built into the design When information such as customer occupation changed from one trade or change in marital status to another, and user needs to know the segregation of transaction between current and previous trade or marital status. The design catered for this process by maintaining a new “snapshot” of the data when change occurred. This was done by attaching a descriptor/time stamp with each occurrence dimension of the data.

TOOLS USED

The presentation layer deployed decision support tools described below. As the Warehouse matures a web interface is being planned.

Oracle Discoverer/2000	This tool will be used for ad-hoc Queries and Reports and accessed ODS/Data Warehouse directly..
Oracle Express Analyzer	This tool was used for MIS needs and primarily accessed Data Marts.

PROJECT PLAN AND SCOPE

The next subsequent stage after a thorough scope study which provided the ingredients for a Data Warehouse development path was to develop a project plan for development of a prototype that will scale into a full fledged Data Warehouse. An 18 week plan was produced to accomplish the following deliverables:

Week	Activity	Output	Involvement
1	User Meetings one to one and workshop	Introduction to DTW Understanding of the business vision Collection of corporate data need	User Departments Business Sponsor I.T Staff
2	One on one/many Interviews	Requirements gathering ROI and return business value	User Departments
	Move Forward Agreement		Business Sponsor
3	Consolidate Information and more interviews Presentation to Users	Meta data collection Gap analysis High level Data Model	User Departments I.T
4	Consolidation of requirements Start of initial prototype Hardware ready	Study infrastructure and define resource needs(Interviews with operations) Gain agreement on business deliverables for MIS Define Strategic Goals, Vision, and Initiatives of the Enterprise Define Objectives and Purpose of Enterprise Data Warehouse Collect Enterprise Information Requirements Create Very High Level Enterprise Data Warehouse Logical Data Model Finalize MIS data mart logical data model Define data warehouse Implementation Road map and schedules	I.T Hardware Vendor End Users Operations

5-8	Detail design	Architecture Document ETT specifications Load of Initial schema	I.T End Users
9-12	Development	Proactive Testing	I.T/End Users
13	Development	Connect to source systems through Prism 80% testing complete for reports	I.T End Users
14	Development	Documentation complete on design, Extraction, load, MIS application etc Data base loaded, reports produced Certification of requirements	
	Checkpoint Reviews		
15	Regression testing and QA business review	Issues, Resolution	All
16	Final tuning and sign off and stress testing	Screens, reports, MIS data mart operational with sub set of live data	User presentation and acceptance before going live
17	Incorporate any changes and further tune the application	Application and systems documentation ready	8,9
18	Pre production with all live interfaces operational	All operational procedures ready and hand over to operations Application changes frozen	All
18	Prototype in Production		
19	Review and move forward		
	Change Requests		

ARCHITECTURE

The following objectives underpin the value of the Data Warehouse to the Bank:

- Provide a common integrated, accurate view of customer and portfolio, services such as Cards ,Deposit Accounts etc
- Support customer -modeling, analysis and target marketing.
- Enable Bank to maximize quality, profitability and size of its customer base.
- Enable result analysis (profitability).
- Enable delivery of market intelligence to customer service points.
- Provide a flexible and adaptive platform for future needs.
- Provide internet enabled interfaces and open technologies.
- As data is cleansed and loaded into the warehouse, provide basis for clean data population mechanism into operational OLTP systems

The following architecture assumptions were made regarding the project:

- The planning and prototyping horizon for the architecture was assumed to be 5 months for prototype and 18 months for EDW delivery
- The scope of the Data Warehouse project takes into consideration 180 day incremental proof of concept and subsequent deployment of the Enterprise Data Warehouse
- After being populated with required data, the Data Warehouse will contained enough information to enable enhancement/replacement of the Customer Data Warehouse with necessary relationships and possible branches into Data marts for data mining needs.

- Users for Data Warehouse include:
 - Market Managers
 - Predictive Modelers
 - Product Managers
 - Finance
 - Campaign Management
 - Executives
 - Branch Managers
 - Initial number of users was estimated at 50.

LOGICAL ARCHITECTURE

Figure 1 below represents the logical view of the architecture. There are several significant points to be noted:

- A hub architecture was used for ETT (Data Extraction, Transformation, and Transport). The alternative would be a point-to-point ETT architecture where source systems interface with the Data Warehouse via point-to-point system interfaces.
- The architecture defines three types of Data Stores:
 1. Data Warehouse for atomic level information
 2. Data marts serving the needs of specific work groups and processes as required
 3. ODS or Operational Data Store to support data access needs associated with operational processes as required.
- A dependent Data Mart architecture was selected. In this type of architecture all Data Marts were sourced via information stored in the Data Warehouse and ODS as required.
 - The Data Warehouse served as a common source of validated and synchronized information and served as the primary source of information for the Data Marts. This avoided the problem where users obtain different answers to the same queries because the information is not consistent across data marts.
 - The Data Warehouse served as a repository of detail and historical information that can be used to augment subset and aggregate information in the Data Marts. The stated goal would be to have the Data Mart handle 80% of the queries with the remaining 20% handled by the Data Warehouse.
 - Having detail historical information stored in a single repository makes it much easier to recast historical information to reflect the results of complex query processing..
 - Meta data activity was curtailed in the interest of time for the proof of concept. However, full blown meta data functionality will be deployed in the complete enterprise Data Warehouse Phase.
 - Extraction process will deploy use of 3GL, Oracle PL/SQL and relational tables and PRISM product set

1

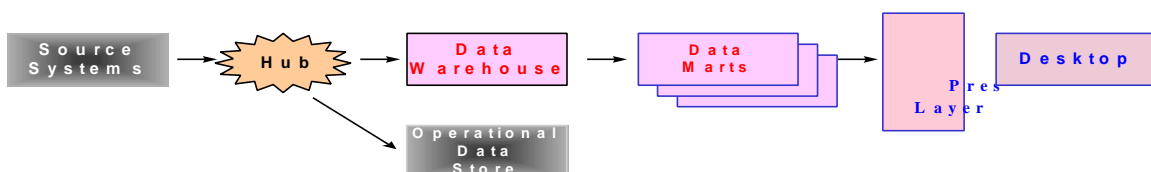


Figure 1

Figure 2 shows another layer of detail in the logical architecture for the long term/strategic Data Warehouse phase. In this layer of the Logical Architecture a number of significant decisions were made:

- The presentation layer accessed the Data Warehouse, as well as, Data Marts.
- The Hub component (Extraction Layer) fed data to the ODS, as well as, the Data Warehouse.

Extraction and Transformation processes extracted data in such a way that data streams from multiple sources were merged or in other ways processed as a group of files. Some of the transformations were relatively straight forward and the data was processed on the fly without staging. Complex transformations were staged into Oracle8.

Figure 2 depicts the underlying Data Ware Architecture:

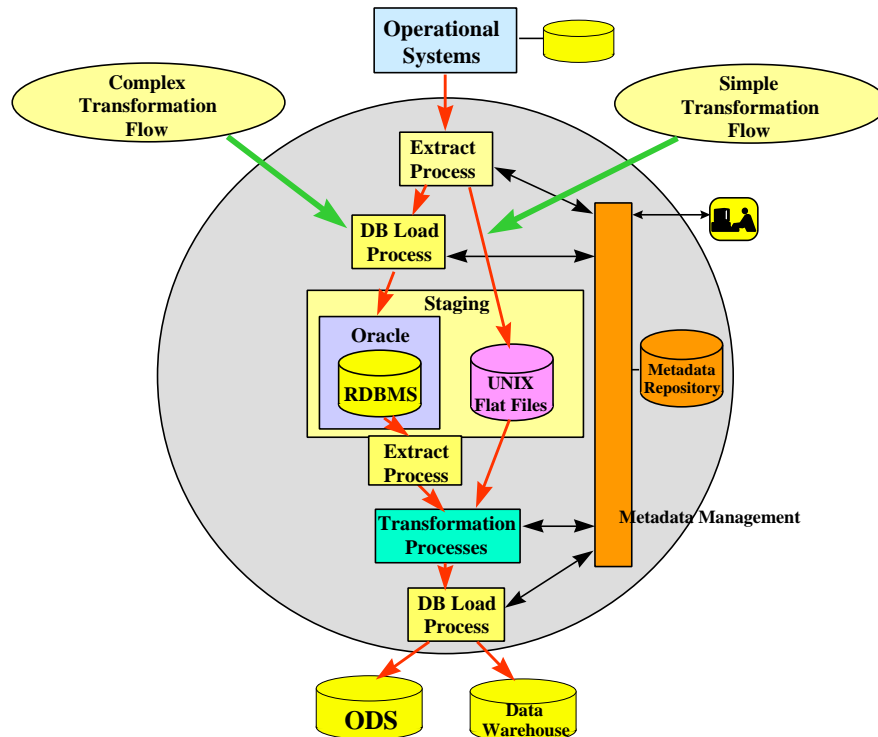


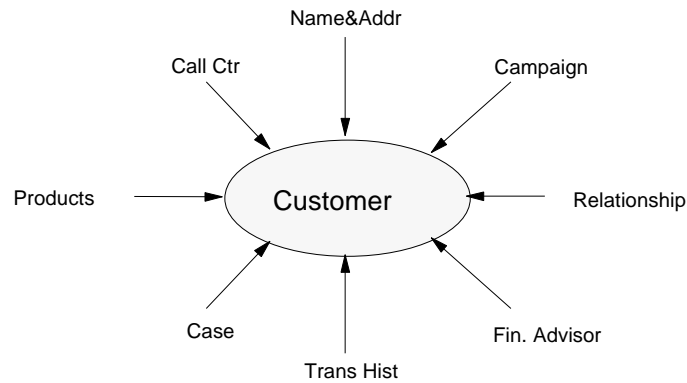
Figure 2

INITIAL DATA MODEL

The data model contained the following key business areas with uni and bi directional relationships.

- customers
- accounts
- products
- branches & regions
- addresses
- banking activities
- Marketing campaign and analysis
- Deposit accounts
- Cash Cards
- Loans
- Transaction history

Sample View of one of the Star Schema



The Warehouse data model started out normalized (i.e. with no duplication of data) and then enhanced for performance and manageability by using:

- derived summaries and aggregations
- controlled data redundancy
- data partitioning

LESSONS LEARNED

This paper summarized the big picture into a very concise description of processes involved in developing a Relationship Banking Data Warehouse. Author may be contacted to discuss more of the project specifics.

Lesson's learned from positioning and delivering the project includes heavy emphasis on an Executive Sponsorship, Awareness of Data Warehousing concepts with End User community and I.T to play a service delivery role rather than forcing business rules and its design on the End Users. A joint RAD(Rapid Applications Development) with active involvement from End User is key to successful Implementation of Data Warehousing projects.