Oracle9i Database for Data Warehousing and Business Intelligence

An Oracle Business White Paper

May 2002
INTRODUCTION

To maintain differentiation from competitors and to provide products and services cost effectively, business managers need to continuously assess relationships with customers and suppliers. More critically, business managers need to repeatedly reassess the answers to three questions every business manager should know - What is your business?, Who is your customer? and What does your customer consider value?

Consolidating data across the organization and deploying business intelligence at all levels is necessary to answer important questions accurately. Many enterprises do not tackle Business Intelligence (BI) due to the difficulties in implementing an infrastructure that can scale and satisfy the breadth complex analytical requirements. This paper describes how Oracle9i Database can assist enterprises successfully deploy a complete and scalable business intelligence system at a lower cost.

BUSINESS INTELLIGENCE PROCESS

A successful BI system begins with consolidating data from transactional or operational systems in order to perform fast, accurate business analysis on high quality data. Poorly integrated or low quality data will deliver poor or worthless information analysis. Extraction, transformation, and loading (ETL) is the process that enables IT systems to consolidate data from multiple sources for subsequent business analysis. In this process, data from transactional systems is extracted, transformed, and loaded into a central repository called a data warehouse.

The data warehouse itself must also be designed for scalability due to the explosion of data growth expected in the next several years. The incredible growth of Internet usage by businesses has also compounded the problem, as website clickstream data is expected to quickly add petabytes to the data warehouse. Additionally, more and more historical data is needed to perform accurate forecasting and analysis. A common pitfall many business intelligence projects experience is not being able to design a data warehouse for growth. Subsequently, customers need to rebuild their data warehouses from the ground up when data volumes overwhelm their original systems.
Once the data warehouse is deployed, it must deliver a high level of performance and maintain fresh data for fast, accurate decision-making. Another measure of success of your business intelligence project is if the data warehouse is really being used. If users executing queries are not satisfied with the response time, they won’t use the system.

With a solid and reliable data warehouse foundation, a business intelligence system can turn high quality data into valuable information in the form of reporting, ad-hoc querying, and sophisticated analysis. Users now have a single source of reliable data to quickly and easily access the information they need to make critical business decisions.

ENSURE BI SUCCESS AT A LOWER COST WITH ORACLE9i DATABASE

Despite having valuable data that can tell enterprises about performance, customer behavior, process efficiency and important trends, many companies do not tackle business intelligence due to high costs, lengthy implementation cycles, and the inability to satisfy a breadth of end-user requirements. Oracle is the only vendor in the market that offers a complete and integrated business intelligence infrastructure for fast deployment at a lower cost. With Oracle9i, businesses can rapidly build data warehouses & data marts, and perform an array of integrated reporting, ad-hoc querying and sophisticated analysis.

With Oracle, you can start off small and grow as needed. Only Oracle9i Database provides a single, integrated business intelligence engine for scalable data warehousing, OLAP, data mining, and ETL. With Oracle9i Application Server, businesses can perform an array of integrated reporting, ad-hoc querying and sophisticated analysis -- all through a portal. With Oracle9i, companies can build out a complete business intelligence system at their own pace for a faster return on investment.

Oracle9i Database goes beyond meeting traditional data warehousing needs for high performance, scalability, manageability, and security by also integrating business intelligence capabilities such as OLAP, data mining, and ETL inside the database itself. No longer do organizations have to move and maintain data in specialized servers for OLAP, data mining, and ETL. With no data movement and a single database engine for data warehousing and business intelligence, companies have a faster path to information at a lower cost.

Effective Data Consolidation Will Make Or Break Your BI Project

Effective consolidation of your data warehouse is the number one success factor of your business intelligence solution. Poorly integrated or low quality data will deliver poor or worthless information analysis. The process of data consolidation or ETL is complex and absorbs most of the budget and time of a data warehousing project.
The reason why data consolidation is so complex is due to the enormous amounts of diverse data sets being generated by different transactional and legacy systems. Non-standardized business terminology needs to be consolidated into a consistent, single view within a data warehouse. For example, a legacy system may enter product information as 'product_id' while a customer relationship management application may enter the same information in as 'item_id'. Whatever the case may be, a new single data model needs to be created within the data warehouse, data then needs to be extracted from these sources, transformed into a single, consistent view, and loaded in the target data warehouse.

Depending on how complex the environment is, data consolidation may require extensive custom coding, training and expertise, additional servers for processing and cleaning data, and ongoing maintenance of this entire process. In order to simplify this process, it is recommended that every BI project use an ETL tool. Typical ETL tools provide built-in connectivity to different data sources, some automatic code generation depending on how complex the transformations and cleansing are, and ETL maintenance. And all of this is enabled through a GUI environment.

ETL tools can dramatically reduce the time required for developing, deploying, and maintaining a data warehouse, but many customers still build their own ETL process. Why would a company go through a very time-consuming and labor-intensive process when ETL tools are available to purchase off the shelf?

**Price, Performance, Personalities**

Price, performance, and personalities are the issues that deter companies from using ETL tools. The price of an ETL tool can be pretty heavy and may not be in the budget even though the price can be justified by increased productivity and long-term savings. Depending on the environment, cleansing and transformations become very complex, especially dealing with huge amounts of data being generated by transactional systems and web sites. Some ETL tools break down and organizations must then rebuild transformations in a procedural language in the database server for example for better performance. The personalities in an organization may be such that they reject anything that will make their jobs simpler. If this is the case, they will bypass the tool and do their own thing. Fortunately, Oracle Warehouse Builder overcomes these issues by providing the most cost-effective, scalable, and easy-to-use solution for designing and deploying enterprise data warehouses, data marts, and e-business intelligence applications.

Oracle Warehouse Builder provides a complete solution requiring almost no coding for extraction, transformation, and loading. This eliminates much of the cost associated with time-consuming custom coding and extensive training. Because the ETL processing engine lives in the Oracle9i database, no separate transformation or staging server is needed which equates to better performance and less cost. The lifecycle management feature dramatically reduces maintenance costs since any change in the source, metadata repository or target is reconciled,
eliminating the labor intensive procedures. For example, if analysis requirements change such as adding another dimension of analysis of a product (for instance: color), you can add this dimension in the metadata repository once and it will automatically extract and load this data. Also, if business requirements change on the transactional system such as adding a new product line, this information will automatically be updated in the data warehouse. This could represent a significant time savings for administrators.

Oracle Warehouse Builder overcomes the concerns about price, performance, and personalities that prevent companies from creating their own ETL solution. First, it is very price competitive. It is free with Oracle Internet Developer Suite. Second, performance is not an issue. Since the ETL engine lives in the Oracle Database, ETL processing is scalable and faster. Third, Oracle users will adopt this tool since they are familiar with the Oracle environment and will embrace the integration with the Oracle database, application server, enterprise manager just to name a few. They will be less likely to grown their own solution and be more productive.

Planning For Performance and Scalability Protects Your Investment

A common pitfall many business intelligence projects experience is not being able to design a data warehouse for growth. E-business is one of the primary sources in the data explosion. For example, web clickstream data is a huge contributor to adding petabytes to the data warehouse. Also, as the number of customer touch points increases, data is increased. Furthermore, escalating end-user demands also play a part, as organizations collect more information and store it for longer periods of time.

Another measure of success of your business intelligence project is if the data warehouse is really being used. If users executing queries are not satisfied with the response time, they won’t use the system. If query results are slow, decisions cannot be made quickly. Large amounts of data, growing number of end-users, and complex analytical processing such as OLAP and data mining all challenge the performance of a business intelligence system.

With Oracle Database's unique performance and scalability features, customers do not need to rebuild their data warehouses from the ground up when data volumes, user demands, and analytical processing overwhelm their original systems. Oracle is the vendor of choice for data warehousing due to the consistent delivery of unique database technologies for improving performance and scalability.

In terms of performance, there are specific unique database features that enhance the performance of data warehouses. These include Materialized Views, Bit Map Join Indexes and Oracle Enterprise Manager.

Materialized Views increase response time to queries by pre summarizing or aggregating data. For example, let’s say there is a Materialized View created for the
total number of wigets sold in the United States in Q1. When a user—let’s say a VP of sales—executes this query with an ad-hoc query tool, the database fetches the pre-computed result which takes less than a second compared to minutes if the data was not pre-aggregated.

**Bitmap Join Indexes** offers an efficient method of indexing especially useful for star queries, commonly used in data warehouses. This feature has been known to improve query performance by a factor of 30.

**Oracle Enterprise Manager** offers automatic memory tuning, proactive performance monitoring, and a Database Resource Manager. Automatic memory tuning chooses the more accurate memory allocation strategy than can be achieved by hand-tuning of memory parameters. Proactive performance monitoring allows performance issues to be identified and fixed ahead of time. Database Resource Manager helps alleviate performance problems by controlling how much database resources is being allocated to each group of users. The resource allocation can be based on a company’s Service Level Agreement (SLA) for example.

In terms of scalability, there are specific database features allow data warehouses to grow into the petabytes. These include portability, Oracle9i Real Application Clusters (RAC), and Oracle9i Partitioning, and parallel query.

Oracle can port to almost any hardware platform. This means the database can run from low end to high end servers. Oracle can scale up to any hardware platform as your data and workload grow.

**Oracle9i RAC** is the only low cost clustering solution in the industry that allows the database to scale.

**Oracle9i Partitioning** offers most flexible partitioning schemes available so that the best combination for manageability and performance is achieved. For example, range partitioning or partitioning by date may offer the best performance in certain warehouse environments whereas list partitioning by discrete values or countries may offer the best performance in another environment. This is a unique Oracle feature since other vendors pre-define data partitions and therefore, you cannot maximize the performance.

Oracle9i Database’s **parallel query** capability subdivides the work for a query across multiple CPUs so that you can not only scale up to virtually any number of CPUs, but also, the query response time is dramatically increased.
Secure Information Access Reduces Risk

Companies providing access to the enterprise data warehouse across all levels of an organization require data security at a very fine level of granularity, often to the level of individual rows in a data table. For example, a VP of Sales should be able to access all sales information while an individual sales representative should only be able to access the data pertaining to his/her customers. Oracle9i Label Security is a unique database technology which allows different levels of users to securely access information from the data warehouse. No longer do business intelligence tools or applications need to build granular access control into their software. Oracle9i Label Security enforces security closest to the data where it cannot be bypassed.

Integrated Analytical Processing Provides Cost Savings

Every business performs some type of analysis. A common example for Online Analytical Processing (OLAP) is sales performance measured by time, geography and customer profile. Sales force, marketing personnel and business analysts need this information to understand and respond to market demands that affect the business bottom-line. A common example for Data Mining is preventing customer attrition, cross-selling to existing customers, acquiring new customers, detecting fraud, and identifying the most profitable customers.

Maintaining separate analytical processing servers that feed from a central data warehouse creates uncontrollable IT costs and leads to inaccurate information analysis. “Data-intensive operations, such as OLAP, ETL, and Data Mining, are best performed closest to the data, within the database, instead of moving large volumes of data to multiple servers” says Richard Winter, President of the analyst firm Winter Corporation. “This approach will reduce integration costs, improve manageability by reducing complexity, and allows for advanced functionality without specialized servers and skills.” Oracle9i Database offers OLAP and data mining from within the database or data warehouse.

Data warehousing practitioners have been able to purchase products with each of these capabilities for years. However, OLAP products typically have their own calculation engine and data storage and data mining products have their own mining engines. In short, the business-intelligence software industry was maintaining at least four ‘data engines’, each requiring its own infrastructure and tools for managing data, its own availability and recovery strategies, its own security mechanisms, and its own parallelism and scalability infrastructure.

With Oracle9i Database, data does not need to move from one specialized server to another when performing data consolidation and complex analysis. With business intelligence processing inside the data warehouse, users have a faster path to information at a lower cost.

Oracle9i Database is the first and only OLAP-ready relational database, supplying efficient resolution of OLAP queries without adding administrative complexity.
Oracle9i OLAP, an option to Oracle9i Enterprise Edition, provides valuable insight into business operations and markets using features previously found only in specialized OLAP databases.

Oracle9i Data Mining, an option to Oracle9i Database, allows companies to build advanced business intelligence applications that mine corporate databases to discover new trends, and integrate those insights into business applications. Oracle9i Data Mining embeds data-mining functionality into the Oracle 9i database, for making classifications, predictions, and associations.

Business intelligence capabilities does not stop with Oracle 9i Database. With Oracle9i Application Server, businesses can perform an array of integrated reporting, ad-hoc querying and sophisticated analysis -- all through a portal.

CONCLUSION

Quick, quality decision-making can contribute to closing deals, retaining valuable customers, cross-selling, and up-selling. Oracle provides a complete and integrated business intelligence system, enabling quick, quality decision-making at a lower cost.

Because Oracle can accommodate growth of a business intelligence system, your investment is well protected. Oracle can successfully scale to manage very large data warehouses running on any platform to meet business requirements in any industry.

Additionally, with business intelligence capabilities such as OLAP, data mining, and ETL inside the database, customers benefit from reduced integration costs, improved manageability, and advanced functionality without specialized servers and skills, and a faster path to information.