

Data Warehousing – Multi tier approach

How it can Benefit your Business Intelligence Implementation

By Agarwal K

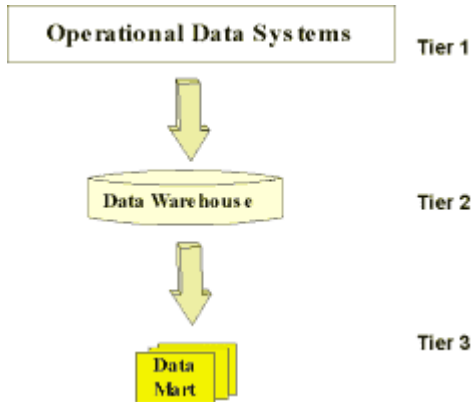
Introduction

This article is intended to describe one of the data warehousing implementation architecture's and its benefits. As with any architecture each business should evaluate what their short and long terms needs are before making architecture and product selections. Specific products usually support specific implementation architectures and some are extendible while others are not. Based on years of experience this three tiered architecture has proven to be one of the quickest to implement while remaining extendible across the enterprise.

Operational data is the source for all business intelligence applications, but the data is typically not in the correct format to support the decision making processes in a business. In fact, one recent study concluded that 93% of all business data was unable to support the business decision making processes. Whether you are implementing data warehouses, data marts, or decision support systems, the goal is always the same - turn this wealth of corporate data into information that can be used in the daily business decision making processes.

What is a Three Tiered Data Warehouse Architecture ?

Three tiered data warehousing means there are 3 tiers of data, each designed to meet a specific set of end user requirements. The picture shows those 3 tiers. The operational data layer is comprised of the data created and maintained by the operational business systems such as order entry, accounts receivable, and general ledger. These systems keep track of specific customers, products, vendors, and resource information on a transaction by transaction basis. They are the systems which support the operations of a business on a day to day basis.



The Second Tier is Called the Data Warehouse. This is where data is moved and cleansed and normalized to support one or many data marts. This data layer may be comprised of multiple data structures; the operational data store (ODS) for tactical decision support applications which require transaction level detail as well as the data warehouse which provides a single common set of data bases designed specifically for all decision support applications in a business. The data warehouse is not designed for direct user interaction, but is designed to provide a single common version of the data elements for all the data marts within the business. This layer is often ignored or forgotten initially and then later added when the expansion of decision support applications includes multiple data marts.

Data Marts are the Third Data Tier. This tier is customized for a specific department or set of users like sales/marketing analysts, financial analysts, customer satisfaction, etc. These users have a specific set of data requirements and want their decision support systems customized so they can access this information to spot trends and take action rapidly.

Everyone seems to understand the need for data marts as these applications are highly effective at meeting end user requirements; what may not be so apparent is the need for the middle data warehouse layer. As was stated earlier, this layer is often ignored or forgotten and this often causes a severe crisis as the decision support

applications expand in breadth or depth within a business. What typically happens is the initial data mart for finance or sales is

successful and expansion occurs: as it does the information staff quickly realizes that the architecture built for a single data mart or two is easily expanded for the 3rd, 4th and 5 data mart applications. In fact, it may take 6 months to build the 1st data mart, 6 for the second, and 6 for the 3rd even if there are common application and data elements which could have been shared, but without the data warehouse infrastructure there is no sharing mechanism to support this requirement. This results in a data warehouse infrastructure that looks like the picture above.

Typically the reason a 3 tiered warehouse infrastructure is not built is because it seems complex, costly, slow and of little value. Ah - but therein lies the fallacy! Implementing this structure initially allows for much simpler expansion and provides a mechanism for data and application sharing and much more rapid implementation of subsequent data marts. Data Mart 1 might take 2-4 months to implement, but data mart 2 may take only 1-3 months, and the third data mart may be less than two months depending on the amount of new application and data requirements.

The common misconception is that the data warehouse must be “fully built” before the initial data mart can be started. This is not true. Incrementally building the data warehouse has proven to be a highly successful method of implementation of both data marts and warehouses, and having the infrastructure in place initially means not having to spend extra time “retro fitting” this layer into an existing data infrastructure.

So What is the Conclusion About 3 Tiered Data Warehouse Architectures ?

Three tiered data warehouse architecture involves taking data from the **operational data bases**, transforming, extracting and cleansing this data into a **single data warehouse environment**, which is then used to build one or more specialized **data mart applications** to support a specific set of end user requirements. This structure can be built almost as quickly as a simple data mart only infrastructure and is more adaptable and expandable as the user requirements grow.