Inmon versus Kimball

W. H. Inmon's approach:

Bill Inmon's approach (top-down) for data warehouse design is an enterprise level, not an approach for departments which have different sets of requirements without considering imposition of standards and integration. He identified need to integrate data from various systems to a centralized repository where the data could be used for strategic decision making.

According to Inmon data should be systemized as subject-oriented, integrated, time-variant and non-volatile structures. The data should be attainable at grain level by drilling down or at summarized level by rolling up. The Inmon data warehouse design pattern is dependent on third normalization form, which can afford opportunity for the granularity of data that provides maximum flexibility to the enterprise. This can prove significantly helpful when organization perspective change for the data to be warehoused. Relationally designed data warehouse can be used to support diverse structure of data OLTP Databases, Exploration Warehouses and Data Mining Warehouses. Design in the top-down approach begins with the consideration of data extraction from the operational data sources. The data is then feed in to the staging area and cleansing is performed. Data is then transformed, integrated and consolidated and then transferred to operational data store (ODS). This data is then loaded into the data marts, which becomes available to end users.

Ralph Kimball's approach:

Kimball introduced the concept of dimensional modeling, which bridges the gap between relational database and multi dimensional databases. He designed the data warehouse in the bottom up fashion by connecting the data marts with a bus structure. The data marts are in

conformance of bus structure in respect of dimension, domain and measure. Data marts are the dimensional model of departmental data which can be viewed as data warehouse (union of data marts) for enterprise level. This arrangement of data marts makes the data warehouse as virtual reality. Therefore, this structure provides the flexibility to place data marts on different servers across the enterprise while data warehouse existence can be considered as virtual (i.e. nothing but sum of all these data marts). Conformed dimension is the backbone of these data marts as these can be assessed more quickly and can share dimensional tables. This approach (bottom-up) is contrary to the top-down approach. In the bottom up approach, data is extracted from the existing legacy systems and then consolidated and verified in staging area. The data is feed in to data store and then more data is added or updated. As the data store contains the fresh copy of data it is integrated and transformed to the data mart structure. The data is than aggregated, summarized and available for end user for analysis and strategic decision making.

Key Differences in approach:

Kimball approach does not indicate the existence of a physically implemented data warehouse. His data warehouse is the union of data marts that are conformed to dimensions. Kimball Bus architecture expresses that he does not recognize a need for a central data warehouse repository. Kimball approach manifests that raw data can be transformed into purposeful information in staging area. Inmon approach is certainly a top-down approach, he consider the dimensional modeling for the design of data marts only and not for the entire data warehouse. Inmon considered the bottom-up approach as brittle, as this design focuses on the end-user requirements therefore it can not reproduce the data in the form which can be prove more purposeful enterprise wise. His approach uses the dependent data mart as the source for star schema usage, solves the

problem of enterprise-wide access to the same data, which can change over time. Inmon emphasized to integrate data from various systems to a centralized repository, his design is at the enterprise level.

Key Similarities/ Agreements in approach:

Both Inmon and Kimball agree that the successful performance or achievement of the data warehouse or data marts depends on effective collection of the business needs. These business needs plays the major role in designing of the data mart which eventually drives the data needed in the data warehouse. They also agree that end-user validation of data mart is significant in order to meet their expectations. Both approaches have the concept of staging, Kimball calls it backroom and Inmon calls it the warehouse. Both Kimball and Inmon approaches recommends extraction of data from a single source. They also share the view that independent data marts or data warehouses can not fulfill the needs of end users on enterprise level for precise, timed and relevant data.

Opinion:

I believe Kimball model is better approach to produce an effective design. Kimball approach is much easier to implement, as we have to deal with small modules to begin with. Also Kimball approach provides more balance in terms of centralized and localized flexibility, as data warehouse is the collection of data marts which can be on different servers across the enterprise. Kimball approach is a quicker and simpler way to create a data warehouse for organization, requires only dimensional structure to fulfill their analysis needs, as in Kimball approach there is no need of designing a data warehouse in 3NF and then going for a dimensional modeling.

References:

Inmon, W.H. (2005). *Building the Data Warehouse*, *Third Edition:* John Wiley & Sons

Kimball, Ralph & Ross, Margy (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*, *Second Edition*: John Wiley & Sons

Ponniah, Paulraj (2001). Data Warehousing Fundamentals: *A comprehensive guide for IT Professionals*: John Wiley & Sons

Breslin, Mary (May 29, 2007) Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models, retrieved on September 12, 2007 from http://www.bi-bestpractices.com/view-articles/4768

Sen, Arun & Sinha P. Atish (2005). A comparison of Data Warehousing Methodologies, retrieved on September 11, 2007 from Communications of ACM

Stephen J. Smith, Alex Berson (December 1, 1997). Components of Data Warehouse, retrieved on September 11, 2007 from http://www.tdan.com/view-articles/4213