Date Warehouse Design Approaches – Inmon versus Kimball

Introduction

We are living in the age where data is of paramount importance to many organizations. In this competitive world, these organizations are realizing that to stay at the top or in some case to survive, they have to effectively extract the information from data. Apart from extracting information, the other problem that they are facing is the amount of data which is growing exponentially year by year. A Few decades ago, the system with data warehouse systems came to the rescue of these organizations. They provide solutions for all aforementioned problems. There are two most prominent approaches to build a data warehouse system: the Inmon architecture and the Kimball architecture. This paper describes each of these architecture styles and attempts to compare the pros and cons of them.

Background

Bill Inmon and Ralph Kimball are two pioneers that started enterprise-wide information gathering, information management and analytics for decision support. These two have produced an immense number of books, articles, seminars in the field of data warehousing. They both support two different schools of thoughts for designing a data warehouse and have near-opposite views on other aspects. The things that are common between them are - they both view the data warehouse as the central repository of data, they both have ETL to load data to a warehouse, they both are used for reporting and analysis purposes. The key differences that they have are the structures and approaches in designing a data warehouse. The question here is that, if an architect starts building a data warehouse, what approach he or she follow? What criterions are there to choose between the Inmon and Kimball model?

The Inmon Approach

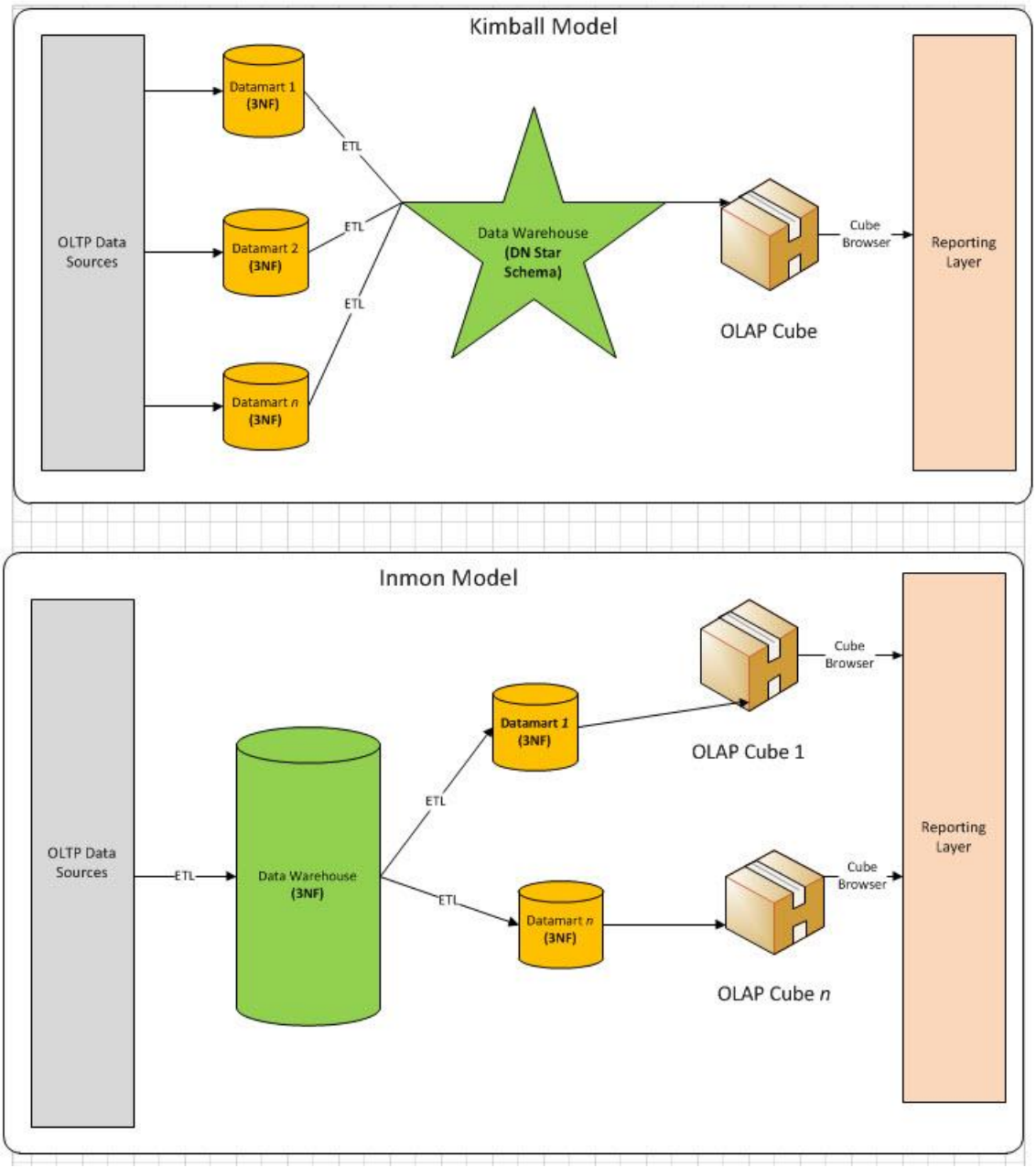
Bill Inmon's philosophy is to create a single enterprise-wide data warehouse for achieving overall business intelligent system followed by several separate databases for each department to serve their analytical and reporting needs, which were later termed as 'data marts'. This approach has received the title of 'top-down approach' as the movement of the data is being done from a more aggregated level to a more detailed level. The entity structure of the data warehouse is built in normalized form.

The key advantages of the Inmon approach are:

- The data warehouse serves as the single source of truth for the entire enterprise.
- Data redundancies are also reduced because of the normalized structure.
- Reduction in data anomalies also make ETL process less prone to failure.
- It is very easy to update the data warehouse as all the things are in one place only.

Some of the disadvantages of the Inmon approach are:

- The structure is more complex and queries can take more time due to large number of tables and joins.
- The initial set-up takes a lot of time.
- More ETL work is involved as the data marts will be built on data ware house.
- Large team of specialist are required to design the architecture.



The Kimball Approach

Ralph Kimball believes in creating several small data marts for achieving department level analysis and reporting, followed by an information bus know as data warehouse. This approach has received the title of 'bottom-up approach' as the movement of data is being done from a more detailed level to a more aggregated level. The dimension model here is not normalized and have to use either a star or a snowflake schema.

The key advantages of Kimball approaches are:

- It is quick to set up and build the project.
- A small team of developers and architects are enough to manage the environment.
- Joining of tables in case of star schema gives a boost in the performance.
- Works really well with department-wise metrics and KPI tracking.

Some of the disadvantages of the Kimball approach are:

- There is no single source of truth involved.
- Redundant data can cause data update anomalies over time.
- Difficult to create reusable structure for different data marts.
- Cannot handle all the enterprise wide reporting need.

References

1. Sakthi Rangarajan. 2016. "Data Ware House Design – Inmon versus Kimball". http://tdan.com/data-warehouse-design-inmon-versus-kimball/20300
2. Benny Austin. 2010. "Kimball and Inmon DW Models". https://bennyaustin.wordpress.com/2010/05/02/kimball-and-inmon-dw-models/