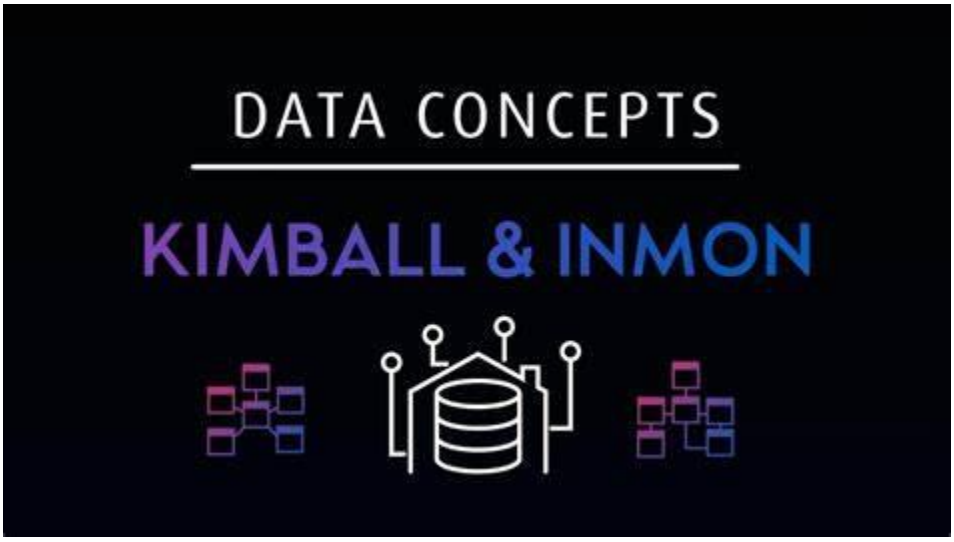
DW Architectures

Aya Mahmoud Mohamed Mahmoud

ITI

Eng. Waleed



Abstract

There are distinctive methods in which we will align distinctive additives of a data warehouse, and those additives are a critical a part of a data warehouse. For example, the data source allows us become aware of in which the records are coming from.

Designing a Data Warehouse is a critical a part of business development. For designing, there are two most common architectures named Kimball and Inmon but the question is which one is better, and which one serves users at low redundancy, so we will evaluate both on some factors.

Introduction

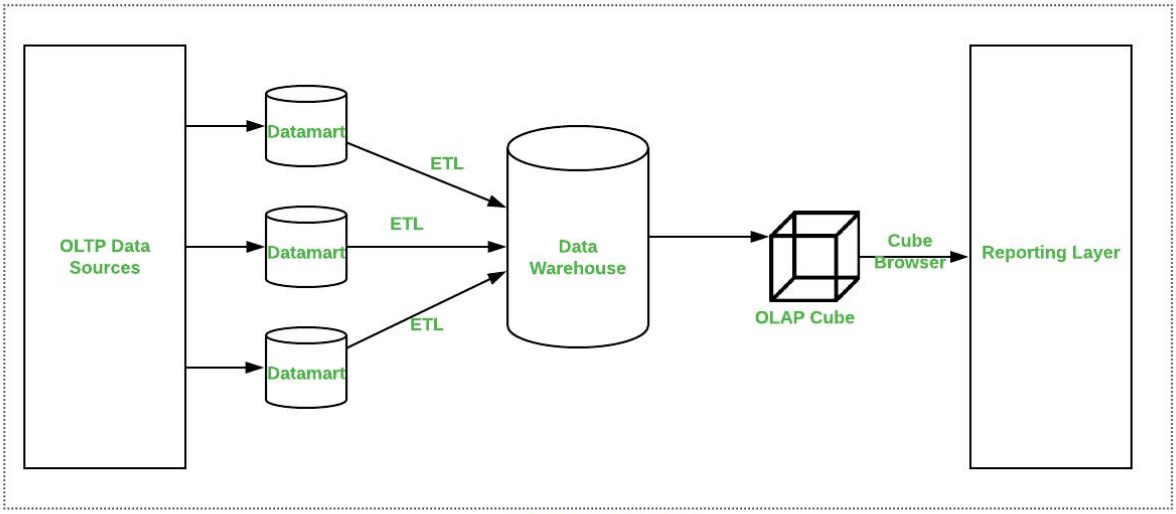
Kimball's approach to designing a Datawarehouse was introduced by Ralph Kimball. This approach starts with recognizing the business process and questions that Datawarehouse has to answer. These sets of information are being analyzed and then documented well. The Extract Transform Load (ETL) software brings all data from multiple data sources called data marts and then is loaded into a common area called staging. Then this is transformed into an OLAP cube. The Kimball model does not have the integration layer. Data moves directly from the source system(s) to the data marts. There is no ODS in Kimball).

Inmon's approach to designing a Datawarehouse was introduced by Bill Inmon. This approach starts with a corporate data model. This model recognizes key areas and also takes care of customers, products, and vendors. This model serves for the creation of a detailed logical model which is used for major operations. Details and models are then used to develop a physical model. This model is normalized and makes data redundancy less. This is a complex model that is difficult to be used for business purposes for which data marts are created and each department is able to use it for their purposes.

Results

For Kimball	For Inmon
Setup and built are quick.	The data warehouse is very flexible to changes.
Generating report against multiple star schema is very successful.	Business processes can be understood very easily.
Database operations are very effective.	Reports can be handled across enterprises.
Occupies less space in the database and management is easy.	ETL process is very less prone to errors.

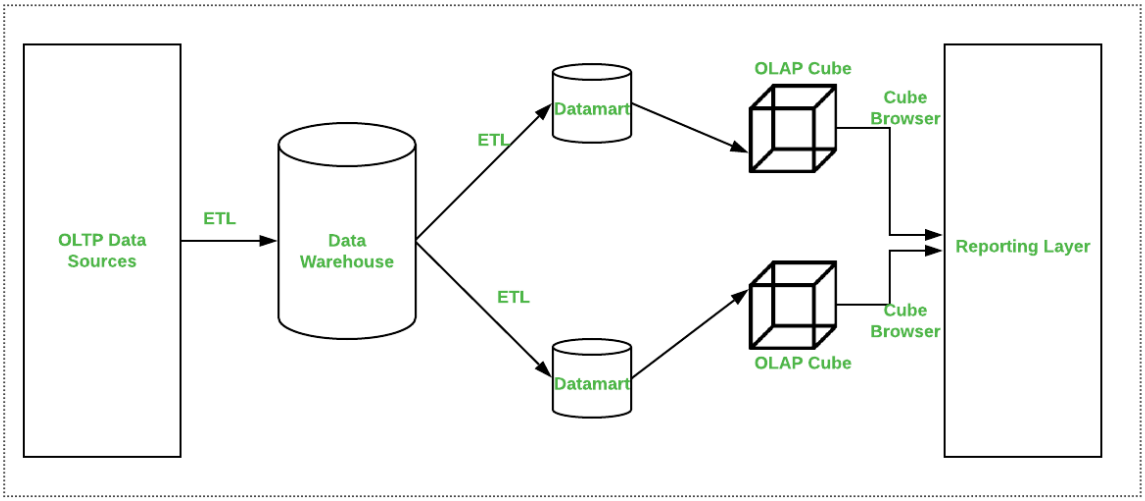
Kimball Datawarehouse architecture:



Kimball Model

Difference between Kimball and Inmon

Inmon Datawarehouse architecture:



Inmon Model

Discussion

The difference Between Kimball and Inmon are as follows:

Parameters	Kimball	Inmon
Introduced by	Introduced by Ralph Kimball.	Introduced by Bill Inmon.
Approach	It has a Bottom-Up Approach for implementation.	It has Top-Down Approach for implementation.
Data Integration	It focuses on Individual business areas.	It focuses on Enterprise-wide areas.
Building Time	It is efficient and takes less time.	It is complex and consumes a lot of time.
Cost	It has iterative steps and is cost-effective.	Initial cost is huge and the development cost is low.
Skills Required	It does not need such skills but a generic team will do the job.	It needs specialized skills to make work.
Maintenance	Here maintenance is difficult.	Here maintenance is easy.
Data Model	It prefers data to be in the De-normalized model.	It prefers data to be in a normalized model.
Data Store Systems	In this, source systems are highly stable.	In this, source systems have a high rate of change.

Which is the best model?

Kimball is the better choice if you want to see results faster, have a small team of engineers, and foresee little changes in the business requirements. Otherwise, the data redundancy could cause anomalies and maintenance costs down the line.

Inmon is the go-to for huge enterprises that wish to see a complete picture of their enterprise data, even if the deployment of the data warehouse is going to cost them more and take longer than Kimball's counterpart.

Conclusion

It has been proven that both the Inmon and Kimball approach work for successfully delivering data warehouses. There are even organizations where a combination of both ('hybrid model') has been implemented. In a hybrid model, the data warehouse is built using the Inmon model, and on top of the integrated data warehouse, the business process-oriented data marts are built using the star schema for reporting. We cannot generalize and say that one approach is better than the other; they both have their advantages and disadvantages, and they both work fine in different scenarios. The architect has to select an approach for the data warehouse depending on the different factors; a few key ones were identified in this paper. Finally, for any approach to be successful, it needs to be carefully thought out, discussed in detail, and designed to satisfy the organization's BI reporting needs and should also gel with the culture of the organization.

Resources

1. Inmon, W. H. 2010. “A TALE OF TWO ARCHITECTURES” InmonCif.com. Accessed May 23, 2016. http://www.inmoncif.com/products/A%20TALE%20OF%20TWO%20ARCHITECTURES.pdf
2. Kimball, Ralph, and Margy Ross. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Third Edition. John Wiley & Sons. 2013. Books24x7. Stanford. 2003.
3. “Data Warehousing Concepts” Stanford.edu. Accessed May 26, 2016.https://web.stanford.edu/dept/itss/docs/oracle/10g/server.101/b10736/concept.htm#i1006297
4. Zentut. 2016. “Ralph Kimball Data Warehouse Architecture” Zentut.com. Accessed May 25, 2016. http://www.zentut.com/data-warehouse/ralph-kimball-data-warehouse-architecture/