INMON vs KIMBALL

1. Introduction:

Almost all businesses and companies in today’s world are driven by data. They consume data, extract it to get information and apply technologies to it to gain knowledge about hidden patterns and trend in that data. It could also be taken to an extent, to say that data is a part of the company’s wealth. This wealth must be handled with utmost care, it must be well organized, easily accessible to the required users. In order to achieve this, data warehousing concepts and modelling techniques are applied to this data. Currently there are two diametrically opposite techniques in use, to design and model data warehouses. One of them was proposed by Bill Inmon who is considered as the father of data warehousing. The other approach was proposed by Ralph Kimball just a few years after Inmon published his approach. Although these approaches follow similar guidelines and have many resemblances, they differ a lot when it comes to structure and modeling, target users, time required and overall costs.

1. Differences:

The first major difference between both the approaches is the designing aspect. While Inmon proposed a top-down design model for a data warehouse, Kimball, just a few years later, proposed a new design methodology, the bottom-up model. Thus, in Inmon’s view, a centralized corporate model is built first. The business processes and dimensions are created whenever they are needed. As opposed to this, Kimball’s designing starts with building a data mart for the most important business process in the organization, then the next, so on and so forth, and at the end combining and aggregating those data marts into a data warehouse. All the other contrasting factors between these two views are due the design methodology used at the start. The Inmon view is like an enterprise wide representation model while the Kimball view is focused on integration of individual processes starting from most essential moving towards lesser important ones. From this, it is evident that we will face data consistency issues when using the Kimball approach as opposed to a strong consistent Inmon model. On the other hand, where Kimball’s model could be generated with a general knowledge of the individual processes simultaneously, the Inmon’s model requires a team of experts and high-level officials to generate in order to achieve an enterprise wide integration. Thus, Inmon’s approach is favorable to an established company with plenty of time and money to spend on a robust model of the entire enterprise, while, a company, for instance, a Start-Up, who has immediate requirement for its business process and not a lot of resources to spend on it can implement Kimball’s approach for quick and easy results. If we compare the complexity of these two methods, Kimball’s method is fairly simple whereas Inmon’s method is quite complex to implement. Another point of distinction between the two can be in terms of data modelling, while Inmon’s model is subject and data driven which can be modelled using traditional Entity-Relationship Diagram, as opposed to Kimball’s model which process oriented and requires dimensional modelling. Maintenance is also one the distinguishing factors between these architectures. There is very little or almost no maintenance issues when it comes to Inmon’s approach as the data is consistent and well-integrated, the operations of roll up, drill down, slice and dice, drill across are easily possible. This cannot be said about Kimball’s model, there can insert, update delete anomalies, as different data marts (processes) may or may not follow the same standards and conventions. Also, when it comes to maintenance, establishing timeframes is very important. While Inmon’s model uses continuous and discrete timeframes, Kimball’s uses concept of slowly changing dimensions (Types 1 through 7).

1. Similarities:

Although both the approaches are considered to be diametrically opposite to each other, they do share some similarities. There are two prominent similarities observed from their models. First is the use of time-stamped data. The time attribute is considered as one of the most important characteristics of a data warehouse. In Inmon’s model, it is referred a time element, whereas in Kimball’s model it is referred as a date dimension. This is so that the end users would be able to query the data at any level of detail such as by day, by week, by month, weekday, holiday, year, quarter. Kimball’s model stores these values, whereas Inmon’s model calculates these values at the time of user query. Now, the second point that these two have in common is in the Data Preparation stage of the data warehousing life-cycle. This is the Extract-Transform-Load (ETL) process. After the data is acquired it must go through a staging phase, where the relevant data is extracted and transformed to follow the standards of the data warehouse (Inmon’s model), or a series of data marts (Kimball’s model), where it will be loaded. Data consistency is required in the both the cases, if two very similar (almost identical) queries return different results in any case, it would damage the credibility of the warehouse or mart in question.

1. Article: Data Warehouse Design – Inmon versus Kimball by Sakhti Rangarajan:

This article is posted on The Data Administration Newsletter. It attempts to contrast and compare the two prominent architecture styles of data warehousing. In my opinion it does a great job of doing exactly that. This paper attempts to answer two major questions in the data warehousing field. First, What style of architecture should we choose for our data warehousing needs? Second, What criteria will help us decide which one to choose? The two approaches are explained briefly along with an example and a model diagram to help us understand these architectures. Furthermore, this article also highlights the major advantages and drawbacks of both the approaches to aid us in our decision-making process regarding architecture choice. Although overall this article looks like a complete package, there are mentions of terms such as conformed dimensions and enterprise bus matrix, which a novice user may or may not understand without proper explanation, which may lead them to choose the incorrect architecture for their warehouse. The author fails to mention her own opinions regarding the issue and states most obvious facts about the two models. She leaves the end decision in the hands of the user trying to be as unbiased as possible.

1. Choice: Kimball’s Design Methodology:

The deciding factors mentioned in the above articles helped me to choose a side on this matter. I favor the bottom-up approach because I believe in building a foundation and then iteratively developing upwards. The process wide scope promises a faster delivery time of the report requirements and is beneficial to use when there is an urgency in project delivery. Although Kimball’s approach may be less flexible, the model is quite simple to implement and requires a relatively small team. Unlike Inmon’s approach, this style takes up very little resources and overall infrastructure and set-up, which is great for start-ups. Also, the ETL process is much simpler than that required in Inmon’s model as it the scope for Kimball’s model is refined for a particular process and not the entire enterprise.

References:

1. Rangarajan, S., V. (2019). “Data Warehouse Design – Inmon versus Kimball.” [online] TDAN.com. Available at: http://tdan.com/data-warehouse-design-inmon-versus-kimball/20300# [Accessed 12 Feb. 2019].
2. Panoply. (2019). “Data Mart vs. Data Warehouse.” [online] Available at: https://panoply.io/data-warehouse-guide/data-mart-vs-data-warehouse/ [Accessed 12 Feb. 2019].
3. Miron, K. (2019). “Inmon v. Kimball.” [online] prezi.com. Available at: https://prezi.com/n5wljckq9sbr/inmon-v-kimball/ [Accessed 13 Feb. 2019].
4. SearchDataManagement. (2019). “How to choose between the Inmon vs. Kimball approach for data warehouse design.” [online] Available at: https://searchdatamanagement.techtarget.com/answer/How-to-choose-between-the-Inmon-vs-Kimball-approach-for-data-warehouse-design [Accessed 12 Feb. 2019].
5. Mor Y. (2019). “Inmon vs. Kimball - The Big Data Warehouse Duel.” [online] Xplenty. Available at: https://www.xplenty.com/blog/inmon-vs-kimball-the-big-data-warehouse-duel [Accessed 15 Feb. 2019].
6. ZenTut. (2019). “Kimball vs. Inmon in Data Warehouse Architecture.” [online] Available at: http://www.zentut.com/data-warehouse/kimball-and-inmon-data-warehouse-architectures/ [Accessed 14 Feb. 2019].
7. George, S. (2019). “Inmon or Kimball: Which approach is suitable for your data warehouse?” [online] ComputerWeekly.com. Available at: https://www.computerweekly.com/tip/Inmon-or-Kimball-Which-approach-is-suitable-for-your-data-warehouse [Accessed 14 Feb. 2019].
8. 1keydata.com. (2019). “Bill Inmon vs. Ralph Kimball”. [online] Available at: https://www.1keydata.com/datawarehousing/inmon-kimball.html [Accessed 15 Feb. 2019].
9. Abramson, I. (2019). [online] Ismll.uni-hildesheim.de. Available at: https://www.ismll.uni-hildesheim.de/lehre/bi-10s/script/Inmon-vs-Kimball.pdf [Accessed 15 Feb. 2019].
10. Dinh, D. (2014). “The differences between Kimball and Inmon approach in designing data-warehouse”. [online] Data is the new oil. Available at: https://technologyinsightscoffee.wordpress.com/2014/06/02/the-differences-between-kimball-and-inmon-approach-in-designing-data-warehouse/ [Accessed 14 Feb. 2019].
11. Breslin, Mary. 2004. “Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models” Business Intelligence Journal, Winter 2004. [Accessed 14 Feb 2019].
12. Vaisman A. and Zimányi E. (Springer 2014). “Data Warehouse Systems: Design and Implementation”. Chapters 2,3,4,5,6. [Accessed 11 Feb 2019].
13. Catherine Ma, David C. Chou, David C. Yen, (2000) "Data warehousing, technology assessment and management", Industrial Management & Data Systems, Vol. 100 Issue: 3, pp.125-135, <https://doi.org/10.1108/02635570010323193>. [Accessed 16 Feb 2019]
14. Golfarelli M, Rizzi S. (McGraw-Hill/Osborne 2009) “Data Warehouse Design: Modern Principles and Methodologies” [Accessed 15 Feb 2019]
15. Prakash N, Prakash D. (Springer 2018), “Data Warehouse Requirements Engineering: A Decision Based Approach” [Accessed 16 Feb 2019]