

# Virtual Data Mart for Measuring Organizational Achievement Using Data Virtualization Technique (KPIVDM)

Ayad Hameed Mousa<sup>a\*</sup>, Norshuhada Shiratuddin<sup>b</sup>, Muhamad Shahbani Abu Bakar<sup>c</sup>

<sup>a</sup>Computer Science, University of Karbala, Iraq

<sup>b</sup>School of Multimedia and Communication Technology

<sup>c</sup>School of Computing, Universiti Utara Malaysia, Malaysia

\*Corresponding author: Ayad.H.mousa@student.uum.edu.my

## Article history

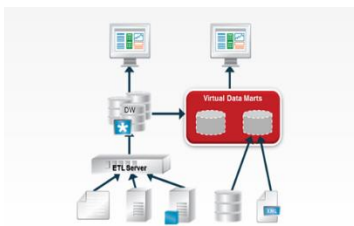
Received :1 January 2014

Received in revised form :

15 February 2014

Accepted :18 March 2014

## Graphical abstract



## Abstract

Currently in the dynamic environment, organizations are confronted with new and growingly vital decisions which can impact their very survival. In fact, these demands are increasing the pressure on Information Technology in order to ensure that data will be delivered properly at the right time and faster rate. In this paper, we propose to build a virtual data mart, especially for Organizational KPIs by using data virtualization technology, which can be used to help KPI developers to build and update performance management system quickly and make these systems work in real time. In this paper, we present a way of identifying and building virtual data marts for Organizational KPIs. The basic principle underlying the proposed approach is that the design of virtual data marts should be driven by the business needs and organizational requirements that each virtual data mart is expected to address. As a consequence, the virtual data mart design process must be based on a deep understanding of the top management's need and users' expectations. A prototype is recommended to validate the use of the proposed method.

**Keywords:** Data integration; data virtualization; data mart; data warehouse; KPI and organizational performance

© 2014 Penerbit UTM Press. All rights reserved.

## 1.0 INTRODUCTION

Currently in the dynamic environment, organizations are confronted with new and growingly vital decisions which can impact their very survival. However, decision-makers who need to measure the performance of their organizations from time to time, have increased their demands on their Business Intelligence (BI) environments to meet their requirements [1, 2]. In fact, these demands are increasing the pressure on Information Technology in order to make sure that the data will be delivered properly at the right time, and quicker than ever before.

The organization faces three major obstacles as on their operations: information inundation, information isolation, and information indecision. Management receives too much information from isolated sources to measure their organization achievement [3-5]. Organizations should measure the performance, efficiency and effectiveness of their processes so as to gain control and competitive advantage. Key performance indicators (KPIs) are used to measure the performance of an organization's processes. Decision makers who belong to any organization depend on the adoption of the entire KPI, and that will lead them for the success of their organizations [2].

The successful future for any organization is determined by what their decision-makers decide, these success will done by know how to accumulate, access and analysis substantial business information which is lead to acquisition effective knowledge , and ultimately make better decisions. KPIs are set of information measures used to evaluate the performance of an organization's processes. KPIs are generated from organizations goals which are quantifiable, measurable, and results oriented [2, 4, 5].

To measure organization achievements,we should provide specific, real time data, and related data, which are considered as significant factors in KPI assessment . Without this, making any decision for the organization will suffer weaknesses such as a lack of precision [4, 5].

Organizations have to be able to build blocks of ad hoc data to be used by performance management systems, where these ad hoc data consists of all information needs of these organizations to achieve their performance. This situation forces top management to pay more attention to data management and how to integrate the right data in the right way at the right time. In this paper, we suggest the use of data virtualization technique to design and develop virtual data mart (KPIVDM) and use it as an input for measuring organizational performance, the KPIVDM can be used to help KPI

developers to build and update performance management system quickly and make these systems work in real time.

## 2.0 DATA INTEGRATION

The process of data integration starts with the extraction of data from heterogeneous or homogenous sources and combine them to make a uniform and multi dimensional interface which

incorporated both form and structure. Its can be referred as data warehouse architecture shows in Figure 1. Using data integration will make it easy for organizations, users, and applications merge different data types ( such as documents, tables and flat files) or different database management system. Data integration creates a base for substantial business needs, defining the relevant data from different data sources and integrate them to generate data warehouse.[5-14].

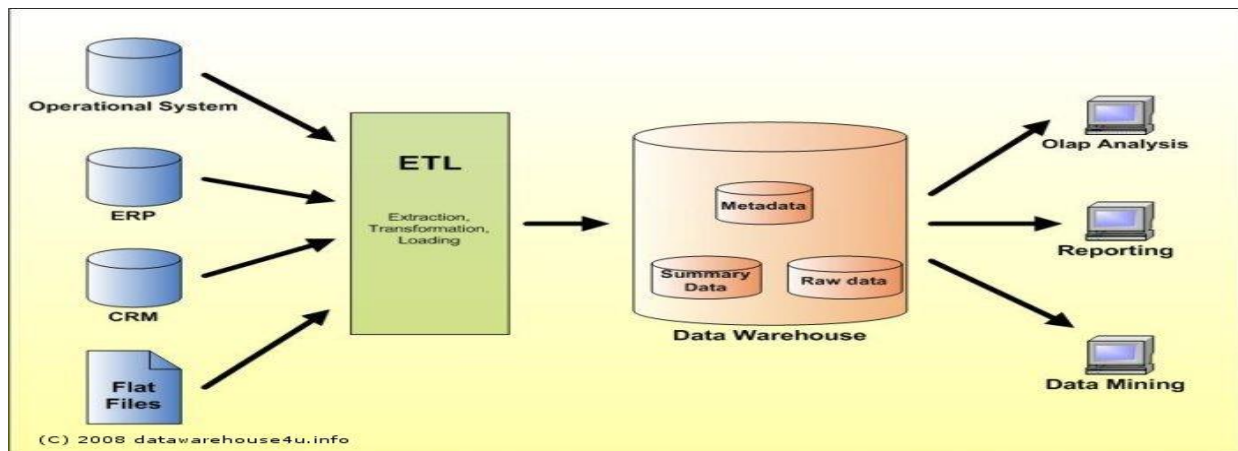


Figure 1 Data warehouse architecture

There are two fundamental concepts of data integration the first one is physical data Integration (data migration) and the second one is logical data integration (data virtualization). In general, physical data integration is the process of building uniform data and build one global schema from different data sources, it is requisite to extract, transfer, and load whole data from heterogeneous data source. The size of data which is migrating from data sources to create data warehouse projects normally would be of immense complexity. On the contrary, logical Data integration (data virtualization) describes the virtual matching of heterogeneous information so as to receive one uniform picture of the data sources [15].

## 3.0 PROPOSED SOLUTION

In this paper, we propose the design and development of virtual data mart which will help software designers and developers to measure the performance of organizations in a perfect and efficient way, and this means it will be to draw the correct data and live data only and prepared to be used as inputs to the process of measuring the performance of the organizations that will be using data virtualization technique, a second concept data integration as mentioned in the previous paragraph

As it is known that the intelligence business depends entirely on the Data Warehouse, but can be considered as the backbone of Business Intelligence. Due to the fact that the Data Warehouse to provide software developers the capability and the ability to handle massive amounts of data. But there remains a problem that is hindering software development: Data Ware House is very expensive and that it suffers from complexity [16-19]. Since most organizations always try to find cheaper economic solutions, data virtualization can be a good option[5, 18, 20-26].

Data virtualization helps to solve this problem by integrating data from multiple underlying BI systems to calculate the required enterprise level KPIs needed. This helps executive to manage the

business at the strategic level without the need to consolidate all the data. Many data integration problems occur for KPIs, especially when the data is stored in multiple sources, have been solved by using data virtualization [18, 20-26].

The purpose of using data mart by designers and software developers is to reduce latency and to provide the real, live and relevant data in the process of measuring performance. Using traditional methods, such as data warehouse will be difficult to get this data. However, the use of virtual data mart which is using a data virtualization technique to provide relevant data that can remove or at least reduce the need to use data Mart, which depend on traditional methods with regard to measuring the performance of organizations [17]. This approach uses abstraction to transform the information in order to meet organizations top management to use this data to help them to create better decisions. See Figure 2.

The organizations have a huge data sources and sometimes the data found in different database management systems ,even though data warehouse and data virtualization technique has solved so many problems and helps the organization's decision-makers to make better decisions, there is still missing specific data that can use for measure organizations performance[20-22, 24]. In order to make the data useable for KPI's Organization appear like a unify form, we propose in this paper to design specific data based on the organization's requirement that can use to create and update organizations performance systems[19-22, 24, 27].

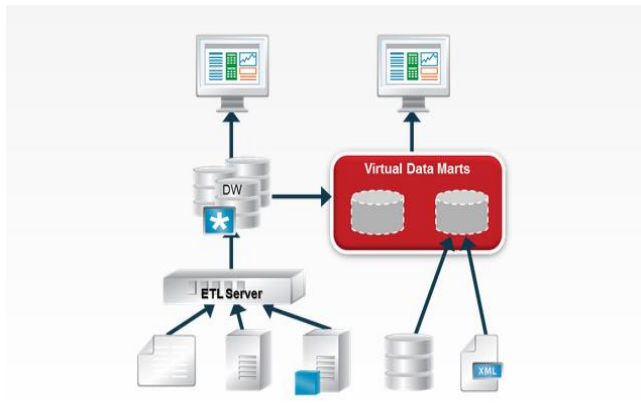


Figure 2 Virtual data mart architecture

Our goal was to have a virtual data mart which is very expressive, and particularly, can consist of all data need for measure organization achievement. The step for develop Virtual Data Mart for measuring organizational consists of various step: i) requirement gathering in the organization (using various method such as interviewing, document review, questionnaire); ii) understand organizations goals and divide main goals to the sub goals to facilitate the goals understanding; iii) identify the information used to measure their organization performance (focus on numeric data); iv) identify data sources from heterogenous source; v) Real Time ETL processes; vi) Build organizational KPIs data. The Virtual Data Mart for KPIs architecture shows in Figure 3.

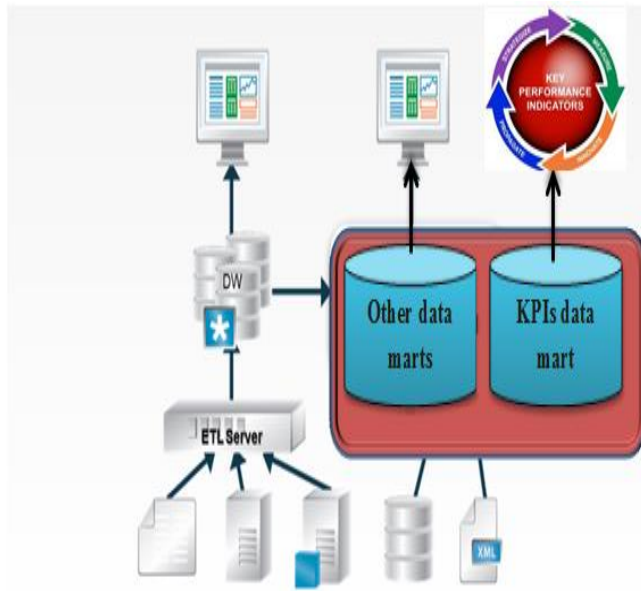


Figure 3 Virtual data mart for KPIs architecture

To sum up, the main benefits of virtual data mart for KPI is that it decreases the access time needed to integrate data , extract it and use it as an input in performance management system (KPIs systems). On the other hand, it increases the response time of queries executed by the data consumers. The ability to provide the information about KPIs by using this technique can be easily established. Moreover, virtual data mart can provide quality data in real time with extract, transform and load (ETL) task. Regarding to

maintenance, our approach will be easy to do during change some or add of data sources, eventually, single version of the related data, even when multiple data sources. Figure 4 will illustrate virtual data mart for KPI benefits.

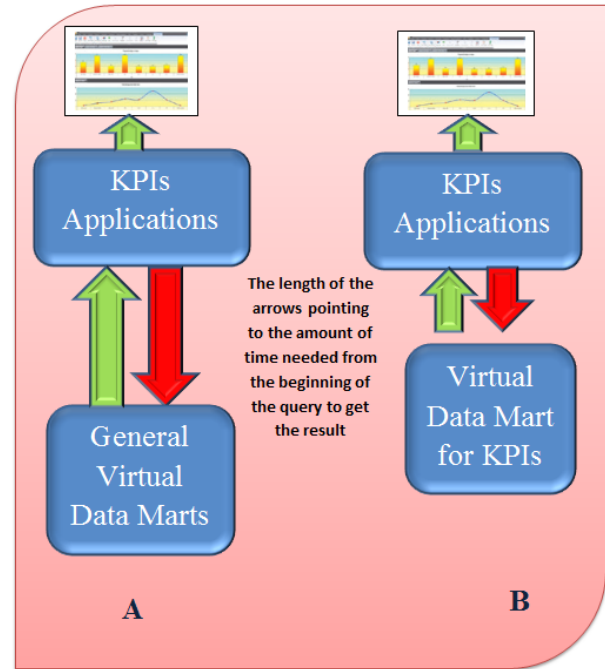


Figure 4 Virtual data mart benefits

#### 4.0 CONCLUSION AND FUTURE WORK

This paper presents identifying and building virtual data marts for KPIs Organizations .The basic principle underlying the proposed approach is that the design of virtual data marts should be driven by the business needs and organization's requirement that each virtual data mart is expected to address. In the future work, we will continue from this research to reach our goals for building comprehensive model that can be help organizations to measure their performance in a proper way.

#### References

- [1] B. Sahay, and J. Ranjan. 2008. Real Time Business Intelligence in Supply Chain Analytics. *Information Management & Computer Security*. 16(1): 28–48.
- [2] C. Ramanigopal, G. Palaniappan, and N. Hemalatha. 2012. Business Intelligence for Infrastructure and Construction Industry. *ZENITH International Journal of Business Economics & Management Research*. 2(6): 71–86.
- [3] S. Jain, K. P. Triantis, and S. Liu. 2011. Manufacturing Performance Measurement and Target Setting: A Data Envelopment Analysis Approach. *European Journal of Operational Research*. 214(3): 616–626.
- [4] X. Ke, W. Li, L. Rui et al. 2010. WCDMA KPI Framework Definition Methods and Applications. In 2010 2nd International Conference on Computer Engineering and Technology. 5.
- [5] V. Masayna, A. Koronios, J. Gao et al. Data Quality and KPI's: A link to be Established. 11–14.
- [6] S. Bergamaschi, S. Castano, and M. Vincini. 1999. Semantic Integration of Semistructured and Structured Data Sources. *ACM Sigmod Record*. 28(1): 54–59.
- [7] M. H. Bateni, L. Golab, M. T. Hajiaghayi et al. 2011. Scheduling to Minimize Staleness and Stretch in Real-time Data Warehouses. *Theory of Computing Systems*. 49(4): 757–780.
- [8] M. Bouzeghoub, F. Fabret, and M. Matulovic. Modeling Data Warehouse Refreshment Process as a Workflow Application.

- [9] M. Castellanos, A. Simitsis, K. Wilkinson *et al.* Automating the Loading of Business Process Data Warehouses. 612–623.
- [10] G. Cavalheiro, A. Dahanayake, and R. J. Welke. Combining Business Activity Monitoring with the Data Warehouse for Event-Context Correlation-Examining the Practical Applicability of this BAM Approach. 263–268.
- [11] B. Devlin, and L. D. Cote. 1996. *Data Warehouse: From Architecture to Implementation*. Addison-Wesley Longman Publishing Co., Inc.,
- [12] J. Eder, G. E. Olivotto, and W. Gruber. 2002. A Data Warehouse for Workflow Logs. *Engineering and Deployment of Cooperative Information Systems*. Springer. 1–15.
- [13] D. Fasel, and K. Shahzad. A Data Warehouse Model for Integrating Fuzzy Concepts in Meta Table Structures. 100–109.
- [14] E. Franconi, and U. Sattler. 1999. *A Data Warehouse Conceptual Data Model for Multidimensional Aggregation: a Preliminary Report*.
- [15] M. Preis, and J. Seitz. 2012. A Hybrid Approach of Data Warehouse Integration Based on New Storage Technologies. *International Journal of Advances in Computing and Management (IJACM)*.1(1): 40–46.
- [16] Munawar, N. Salim, and R. Ibrahim. 2011. Towards Data Quality into the Data Warehouse Development. In 2011 Ninth IEEE International Conference on Dependable, Autonomic and Secure Computing. 8.
- [17] J. Nasina, and S. Kondabolu. 2011. Creating a Virtual Data Warehouse for Manufacturing Industry. *The IUP Journal of Operations Management*. 10(2): 47–59.
- [18] J. Nasir, and M. K. Shahzad. 2007. Architecture for Virtualization in Data Warehouse. *Innovations and Advanced Techniques in Computer and Information Sciences and Engineering*. Springer. 243–248.
- [19] H. Plattner, and A. Zeier. 2011. *In-memory Data Management: An Inflection Point for Enterprise Applications*. Springer.
- [20] L. Weng, G. Agrawal, U. Catalyurek *et al.* An Approach for Automatic Data Virtualization. 24–33.
- [21] R. Eve, and J. R. Davis. 2011. *Data Virtualization: Going Beyond Traditional Data Integration to Achieve Business Agility*: Composite Software.
- [22] M. Ferguson. 2011. *Succeeding with Data Virtualization High Value Use Cases for Analytical Data Services*.
- [23] B. Hopkins. 2011. Data Virtualization Reaches the Critical Mass.
- [24] R. F. v. d. Lans. 2012. *Data Virtualization for Business Intelligence Systems: Revolutionizing Data Integration for Data Warehouses*.
- [25] J. Richter, L. McFarland, and C. Bredfeldt. 2012. CB4-03: An Eye on the Future: A Review of Data Virtualization Techniques to Improve Research Analytics. *Clinical medicine & research* 10(3): 166–166.
- [26] R. Van der Lans. 2012. *Data Virtualization for Business Intelligence Systems: Revolutionizing Data Integration for Data Warehouses*. Morgan Kaufmann,
- [27] J. Hammer, M. Schneider, and T. Sellis. Data warehousing at the Crossroads.