ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

Overseeing Warehouse Design Using MDA in Entire Lifecycle

Ziyati Elhoussaine, Ouzzif Mohammed and Mounir Rifi
RITM laboratory, Computer science and Networks team
ENSEM - ESTC - UH2C, Casablanca, Morocco

Abstract*—Model Driven Architecture (MDA) [5] is presented as an approach to system specification and interoperability, defined by the use of several formal models in order to describe the complete lifecycle of designing, deploying, integrating and managing applications. Beside, many conceptual methods have been exposed to solve different parts of a Data Warehouse (DW) system. However, they do not cover all the process, they are a partial solutions dealing with subset of DW aspects and do not offer designers an integrated method to oversee different stages of DW (Data Source, ETL process, DW repository and so on). In this paper, we explain briefly the dependently concepts of Data Warehouse and propose a model to check and control properties during data warehouse design based on generic CIM, PIM and PSM

***Index Terms*—MDA; Data Warehouse; Availability; security; maintainability; dependability; reliability.**

I. INTRODUCTION

The architecture of a DW is usually depicted as a multi-tier system in which data from one layer is derived from data of the previous layer [2] as presented in Figure. 1.

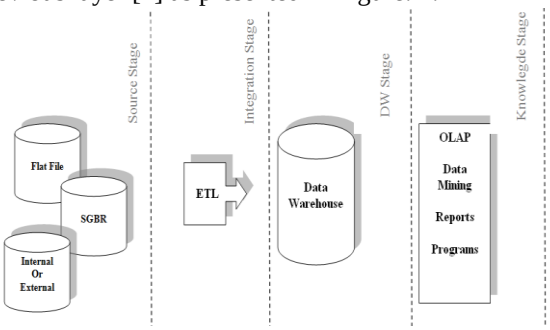

Fig. 1 Data Warehouse Architecture

Such proposition has been presented for solving such a level of DW (ETL process or dimensional modeling (MD)). Interoperability and integration may produce additional problems. [3] Try to deal these problems by proposing a method based on UML (Unified Modeling language) [4] and the UP (Unified Process).

MDA [5] looks ideal for this task to design the complete lifecycle of developing applications by using models in software development; it is a standard for defining requirements, and non-functional conditions that should be taken in consideration, such as reliability, security and maintainability.

So the goal of this paper is to verify that the implementation worked as desired in term of a set of requirements and specifications.

This paper is organized as follows: sections 2 will present the related works on the development of dependable DWS; section 3 will give an overview of dependability aspects and MDA; section 4 introduces our approach applied to dependability of DWS. An example of implementation is shown in section 5, and finally in section 6 we present a conclusion and some future work.

II. RELATED WORKS

Briefly, Model Driven Approach Figure. 2 is a method for software design based on three stages:

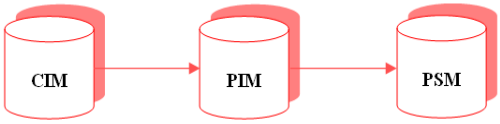

Fig. 2 Generic MDA models

CIM: Computation independent model, approach helps to focus on essential part of the solution designed, separated from platform details [12].

PIM: Platform independent Model provides formal specification from system structure and function that is free from any platform [11]-[10].

PSM explains how particular technology can be used to implement the function described in PIM [9].

There are many works implying MDA in DW design, [6]-[7] align the design of Data warehouse with MDA process, the model Driven data warehouse, this is for one level (ETL e.g.) or all levels.

Also [13] define a framework using MDA, the main idea at this point is once the platform independent specifications are developed the rest of models from them can automatically derived by corresponding transformations, they define set of non-functional – requirement for design process. This interaction between MDA and DW is benefit, in the end of process, DW schema is generated, others benefits of this junction are:

Business perspective: focus on developing conceptual models.

Portability: adopted for different technologies.

Adaptability: since PIM as described above is Independent platform.

In other hands, computer dependability [14] a set of concepts involving reliability, availability, security and maintainability. Capturing non-functional requirements without implying them into the conceptual model for

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

different levels may cause problems, so it is recommended to involve and measure theses dependability components throughout the phases of data warehouse design based on their importance.

As an example for this requirement Security; security policy known to ensure authorization on the DW scheme. Authors in [16], defines a model which propose a security concept for OLAP. According to these security rules, a restriction of data is related to each role. In [17] authors show how security policy for both stages DW and OLAP focus specifically on expressiveness and usability, but only focus on acquisition, storage and on the front-end side. None of them examine the representation of security at the early stages of the DW design.

Other proposition [18] a security design methodology similar to the classical database methodology (requirement analysis, conceptual, logical and physical design) which covers requirements and concrete implementations in commercial systems. In addition to security many aspects can be treated in the same way, and satisfy all requirements for different layer of conception and cover all stages.

III. Proposal Model

During DW conception we imply a set of dependability factor [14] in order to design a consistent architecture and satisfy to all requirements and specifications imposed earlier. Figure. 3 resume our approach, involving at each stage of design a set of defined parameter for MDA generic.

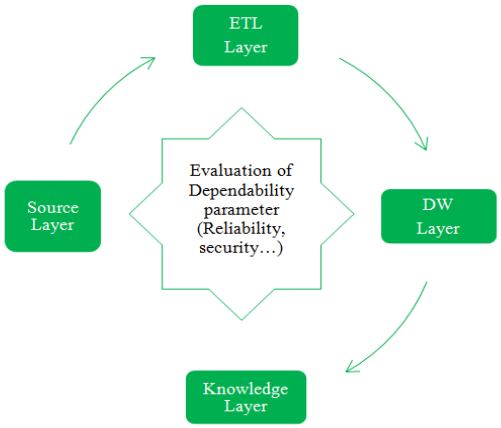

Fig. 3 Proposal model

The framework will deal with three dimensions stages, dependability and finally timestamp. Figure. 4 depict these dimensions with hierarchies.

The cost of dependability is not covered at this stage, a simplified cost is adopted by considering exist or not during a process or the percentage importance.

As shown in Figure. 5, Dependability is considered as a dimension where we define a hierarchy (CIM>PIM>PSM Higher level to lower level) of aggregation, for dimensional Modeling.

As an example for the data in MDA (dependability) dimension is shown as follow in figure 5:

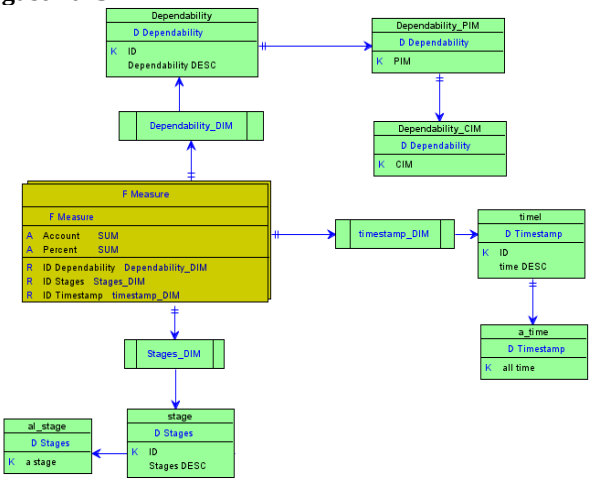

Fig. 4 Dimensional schema of our framework

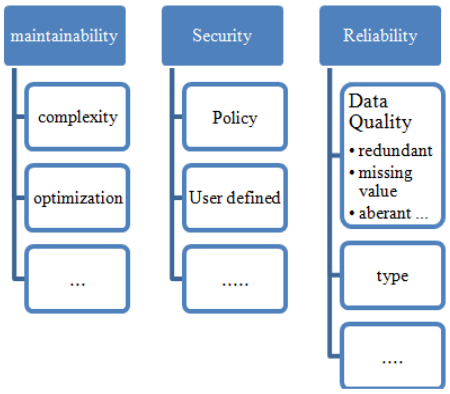

Fig. 5 Snapshot of defined attributes for CIM and PIM

Another presentation is presented in [14] where authors use SIG (Soft goal interdependency Graph) and UML for describing this requirements Maintainability, Reliability & Availability and Security. This approach had the ability to follow this measures defined above in time (during design time).

IV. Experience

To validate our model Data warehouse design control, we implement the framework in Oracle 11g2 [14], and then we report many aspects of measurement across conception stages.

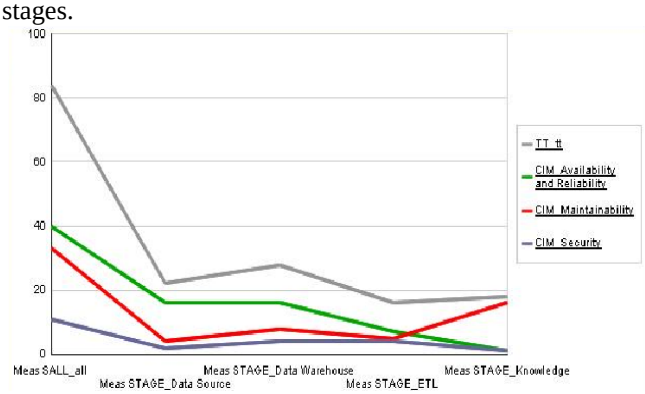

Fig. 6 Some aspects behavior throughout DW stages

Figure. 6 gives a global vision for the whole process of

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

design by considering Reliability, maintainability and security throughout the project design (Data Source, DW, ETL…), thus, a comprehensive view of the behavior of the rules imposed initially.

Drill down to specify a dependant factor to see its impact and behavior for all stages during time will give other additional information to designer for balancing requirements Figure. 7.

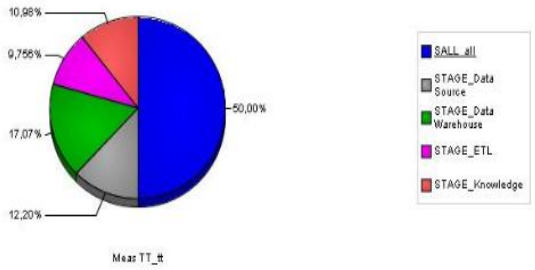

Fig. 7 Example of specified requirement's measures for all DW stages

For a given MDA parameter, we can easily know which related factor influent more than other; this is ensured by the framework and offer remarquable performance with navigation by choosing the adequate level, Figure. 8 give an example for maintainability and its related components.

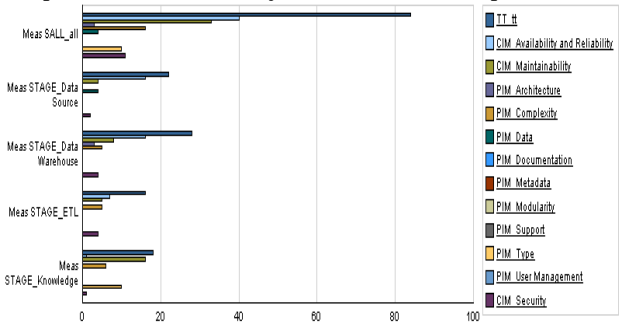

Fig. 8 Drill down operation result

V. CONCLUSION

In this paper, we presented some research works which focused on the DW design using MDA approach. Those works were proposed for limited parts of DW. In fact, we present a framework schema to imply and oversee different requirements involved for designers during DW lifecycle. Experiments are then exposed to prove the usefulness of our method, thus, administrator can verify at any stage of DW process the correctness of his proposition overlooked to the earlier specification. As future work, we propose to enhance our framework by dealing with an effective cost of each specification such as (security, availability…).

This work gives the opportunity to measure the accuracy to a designed data warehouse, up to now directly adopted after design for a given enterprise.

ACKNOWLEDGMENT

This work has been supported by RITM laboratory research group unity in association with ENSEM institute (National School of Electricity and Mechanics), Support for the carrying out doctoral study program's and support for the development of doctoral studies at Hassan II, Ain chock University.

REFERENCES

[1] Jacobson I., Booch G., Rumbaugh J. The Unified Software Development Process. Object Technology Series. Addison-Wesley. 1999.

[2] Jarke, M., Lenzerini, M., Vassiliou, Y., Vassiliadis, P. Fundamentals of Data Warehouses. Springer. 2000.

[3] Luján-Mora S., Trujillo J. A Data Warehouse Engineering Process. In Proceedings of the Conference on Advances in Information Systems (ADVIS'04). Lecture Notes in Computer Science, Izmir, Turkey, October 2004. Springer-Verlag.

[4] Object Management Group Language Specification 1.5. http://www.omg.org/cgi- bin/doc?formal/03-03-01

[5] J. Miller and J. Mukerji, "MDA guide version 1.0.1, "2003".

[6] Poole J. Model Driven Data Warehouse (MDDW). www.cwmforum.org/POOLEIntegrate2003.pdf, 2003.

[7] J.Mazón, J.Trujillo, An MDA approach for the development of data warehouses, An MDA approach for the development of data warehouses, 1st issue, Vol. 45, Elsevier Science Publishers, 2008.

[8] OMG Common Warehouse Metamodel (CWM) Specification 1.0.1. http://www.omg.org/cgi-bin/doc?formal/03-03-02, 2002.

[9] Almeida, J.P.A., 2006, Model Applications, Ph.D. Thesis, Centre for Telemetric and Information Technology, University of Twenty, Netherlands.

[10] Kim, H., 2008, Modeling of Distributed SyMDA, IAENG International Journal of Computer Science, 35:4, 20 November 2008

[11] Vidales, M.A.S., García1, A.M.F., and Aguilar, L.J., 2008, A new MDA approach based on BPM and SOA to improve software development process, Tékhne, 2008, Vol VI, ISSN: 1645-9911.

[12] Pokraev, S.V., 2009, Model-Service-Oriented Applications, Ph.D. Thesis, Centre for Telemetric and Information Technology, University of Twente, Netherlands

[13] J.-N. Mazón, J. Trujillo, M.A. Serrano, M. Piattini," Applying MDA to the development of data warehouses" In proceeding of DOLAP 2005, ACM 8th International Workshop on Data Warehousing and OLAP, Bremen, Germany, November 4-5, 2005, Proceedings

[14] I. Hilal, R. Hilali, M. Ouzzif Modeling Dependability of Data Warehouse System". (IJCSIS) International Journal of Computer Science and Information Security, Vol. 11, No. 6, June 2013, pp 47-54.

[15] www.oracle.com/fr/technologies/datawarehousing/

[16] R. Kirkgöze, N. Katic, M. Stolba, A.M. Tjoa, A security concept for OLAP, in: Proceedings of the 8th International Workshop on Database and Expert System Applications (DEXA'97), Toulouse, France, 1997, pp. 619–626.

[17] W. Essmayr, E. Weippl, F. Lichtenberger, W. Winiwarter, O. Mangisengi, An authorization model for Data Warehouses and OLAP, in: Workshop on Security in Distributed Data

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 2, August 2013

Warehousing, in Conjunction with 20th IEEE Symposium on Reliable Distributed Systems (SRDS'2001), USA, 2001, pp. 9–13.

[18] T. Priebe, G. Pernul, Towards OLAP security design – survey and research issues, in: Proceedings of the 3rd ACM International Workshop on Data Warehousing and OLAP (DOLAP'00), VA, USA, 2000, pp. 33–40.

AUTHOR BIOGRAPHY

Dr Ziyati Elhoussaine is a Professor, in Computer Engineering department in ESTC Institute of Technology, Casablanca, Morocco, and received PHD degree in Computer Science from Mohammed V. University in 2010, presently, his domain area is Business Intelligence, BigData Networking and Data warehousing.

Dr Ouzzif Mohammed is a professor in the computer science department of the higher school of technology of Casablanca - Hassan II university of Morocco.

.

Dr Rifi Mounir is a professor in the computer science department of the higher school of technology of Casablanca - Hassan II university of Morocco; actually he is the Director of the ESTC.