

Model Analysis of Data Integration of Enterprises and E-Commerce Based on ODS

Zhigang Li, Yan Huang and Shifeng Wan

College of Information Management, Chengdu University of Technology,
Chengdu 610059, Sichuan, P.R. China

cdlglzg@163.com cherrishhy@163.com wanshifeng2004@21cn.com

Abstract. Based on the discussion of ODS (Operational Data Store), enterprise data environment features and data integration requirement, three granularity level data models of Web logs in e-commerce are analyzed. Further, data updating and collecting train of thought has been expressed which is realized by granularity manager. Finally, a model is put up, which integrates e-commerce data and other enterprises' data by ODS.

Keywords: *ODS, Electronic commerce (E-commerce), Enterprise data, Integration*

1. INTRODUCTION

Electronic commerce systems are the ones that are built on the Internet and other networks to realize enterprise's business objectives. They can meet the needs of production, sales and services and also can support enterprise's collaborations with its business partners. The electronic commerce systems improve level of enterprises' information automation, management and decision-making. In addition, they offer enterprises a commercial intelligent computing system. As we know, E-commerce brings about enormous competition and business opportunities for the global economy, so e-commerce systems become global information systems. With increasing needs of e-commerce and pressure from enterprise internal development and external competition and collaboration, in enterprises, the demands for information system support are ever increasing as well. Information became the critical resource to enterprise's survival and development. So E-commerce systems are facing big challenges, which are evidenced in that they are required to be open, capable of connecting with other various application systems to integrate enterprises' information resources and improving enterprise's ability of developing markets. However, the more the layers of enterprise's system are, the more difficult data integration is. Thus, Operational Data Store is a better option to solve the data integration problems of e-commerce data and enterprises' other systems data [1, 2].

2. THE REQUIREMENTS OF ENTERPRISES' DATA ENVIRONMENT AND DATA INTEGRATION

Because of the differences among enterprises' sizes, types and industries, data environment varies from different enterprises, so does the requirement of data environment. Data environment characteristics of big enterprise are as follows: firstly, on the whole, there is an operational data environment, and analysis data environment is at the rudimental level, then long-term data has been accumulated, finally, data are relatively canonical. While small and Medium-sized enterprises' characteristics are: firstly, some have operational data environment, but lack the qualifications of building large-scale analysis data environment, then they don't own accumulating long-term data, finally they lack unified data canonical [2].

Enterprise information system is complicated and its information processing contains several courses which are information collection, information management and information controlling. What's more, in connection with different levels and applied aims, it processes information at different abstract level. Then we may find that some terms which are used to describe the processing course, such as data processing, information processing and knowledge processing. At the same time, these terms also reflect the complexity of information processing. With development of E-commerce systems, information system integration is not only the integration of different data information, but also will become integration of knowledge.

The demands of information determine the demands of data integration function. And the purpose of enterprise data integration is to effectively integrate information which distributes in autonomic and isomorous partial data source, so as to realize sharing information among subsystems. Meanwhile, information integration also needs to solve the problems of information, knowledge (experience included) and transformation among data. Because different enterprise data distribute in different business systems, a unified platform is needed to show the analysis results. But this platform is lacking, so a unified enterprise data view can't be offered to the decision-makers. In this case, we need a platform to show the analysis results. We may find that it breaks the barriers among systems, then integrates management information and business information, and offers a unified enterprise production data view and operation data view for the decision-makers and executive officers. Next, some requirements are also needed which are data consistence, data security and efficient use of data [3]. Finally, the ultimate goal of data integration application is to realize enterprise decision-making supporting, which is the real significance.

3. DATA CHARACTERISTICS IN ODS

ODS is a kind of store technology, lying between Data Base (DB) and Data Warehouse (DW). Comparing to Data Base, the way of organizing data for ODS also faces to subjects and integration, just like Data Warehouse. The structure of ODS is mixed, supporting operational transaction process and analysis process. ODS data is integrated, variable, specific and the current data or approaching current data and

faces to subjects [4]. In addition, when running a system, ODS is the place for enterprise to release information, and the information is real time and about to real time. Meanwhile, other systems of the enterprise can use the information also, including Data Warehouse. But there still are some differences between ODS and Data Warehouse. Firstly, their data magnitude varies greatly. In terms of storing time, DW contains lots of historical data, whose amount greatly exceeds the current data and the latest data of ODS. And then their applying demands are different. DW is used for long-term strategy analysis, which mainly faces to professional data analyst and top management [5, 6]. But ODS is mainly used to global OLTP and the latest LOAP, which faces to middle management for daily management and short-term decision.

When processing lots of historical data, Data Warehouse exerts great efficient, but it can't do well to some old systems. Due to various reasons, these old systems lack synthesis, which makes it difficult to analyze data, dispose data and supply data to Data Warehouses. However, ODS can deal with the data offered by old systems better, in a short time and at lower costs, through redesigning data and data processing model [3].

Moreover, an important online technology is also needed, which can actively offer our customers with their interested subject information to gain more potential customers. According to customer's individual needs, this technology also can automatically search data, collect data, filter data and adjust data at the background, and finally supply customers with their needed information, which reduces customers' works in online searching, comparison and negotiation. These high-powered operations are too hard to proceed in Data Warehouses. But in ODS structure, these can be easily done. Once the incorporative and refined results which coming from enterprise's Data Warehouses are stored in ODS, they can benefit e-commerce web environment. Thus, in the web environment, ODS is one of the key structures for enterprise data processing, data in the web environment can be mostly withdrawn from ODS [7].

4. INTEGRATION MODEL FOR WEB DATA AND OTHER ENTERPRISES' DATA

Data model can help users to review how diversified data are integrated and help users to comprehend the final results. When the products are delivered, Data model also ensures the same expectation between the model builder and the end users. The use of data model helps to reduce the originating rate of redundancy. For data model can make redundancy obvious and delete it. So we may see data model is helpful to reduce the overall risk of the program. If data model were not used, e-commerce workmen will find it is very difficult to withdraw data from web daily and integrate them with other enterprises' date, for more interfaces are required to startup the circular developing process. Without the support of data model, it will be a big challenge for enterprises if they want to manage enterprise data as the critical resource [7].

In order to analyze the enormous data from daily web and then integrate them with other excessive enterprises' data, data granularities at different levels must be disposed. In e-commerce environment, daily web data are at the lowest level of granularity. In a day, the daily data can be taken from the web at a certain length of time [8] [9]. The daily web data can be directly introduced into the DBMS of the ODS in the website, which is the level 0. The introduction process of web log to the DBMS can be designed, or can be completed by the practical tools. From level 0, the daily data and enterprises' data are integrated in the ODS, which is level 1. For example, the customer information from the web log is integrated with the data in the enterprises' ODS. Finally, these data from one day or several days are integrated in ODS is the level 2, which can be used for quick report access [7] [8]. All these three levels of granularity need data models.

Web site data need to integrate with other enterprises' data. To support the integration, many data models will be used, every one of which plays a very important role. These models include [7]: Subject Model, Enterprise Logic Data Model, data models of Data Warehouse and data mart, ODS Model of web sites and Enterprise ODS model.

Subject Model—used to better understand all areas, such as product classification area.

Enterprise Logic Data Model (Enterprise Data Model)—used to connect different subjects, such as the connection between customers and products.

Data Warehouse Model and Data Mart Model-- used to meet the purposeful operational demands, such as doing sales report forms.

Website ODS Model—used to take use of website log's path information.

Enterprise ODS Model-- used to integrate web data and other enterprises' data.

Next this paper will mainly introduce the Enterprise ODS Model:

The structure of Enterprise ODS Mode is integrated and subject-orientated. What makes it different from Data Warehouse is that enterprise ODS is timeliness, including limited historical data. In the e-commerce environment, enterprise ODS is regarded as the Level 1 granularity. Enterprise ODS is renewed from the source systems (Web site ODS and other business systems), and other business operational systems or data mart [6, 8]. In other words, the enterprise ODS is the combination of all kinds of ODS.

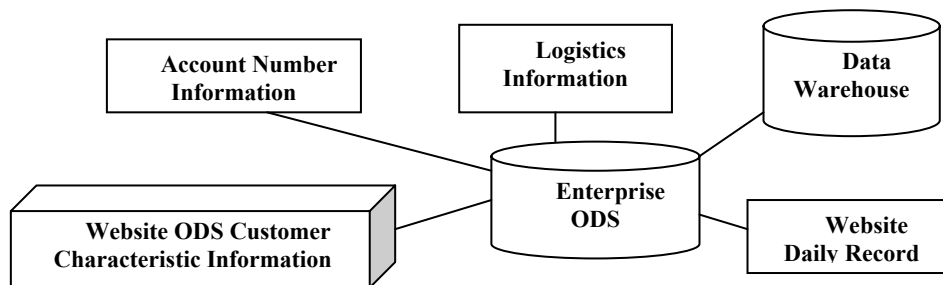


Figure 1. Data Flow Mode

Enterprise ODS environment may be very complicated. According to WH Inmon’s analysis and giving consideration to the factual situation of enterprise’s website [7, 9], we can design a data flow model. From this model we can see all the data flows from the source systems, Data Warehouses and data marts

Enterprise ODS supplies data to website ODS, so it can make some processes run together, which can take advantages of the parallel structure of this platforms and ETL tools. Internal customer entity of the enterprise ODS plays a very important role in the e-commerce environment. Figure 2 shows some of the data models established for customer entity in enterprise ODS.

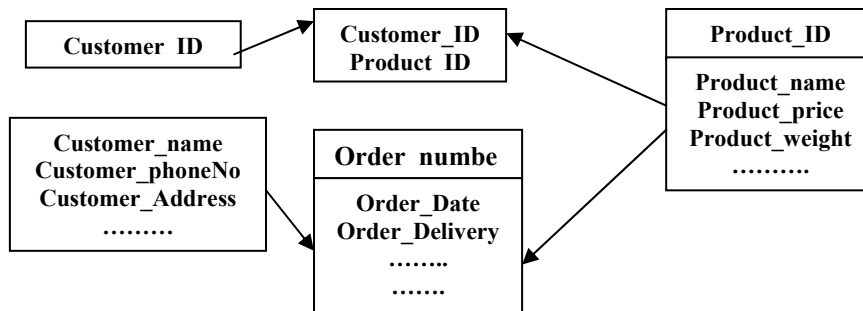


Figure 2. Data Model of Customer Entity in the Enterprise ODS

Using the Granularity Manager, ODS can fit different levels of data granularity (Level 0, Level 1, Level 2). Through using the three levels in web environment, we can successfully implement data updating and summary. And the levels of granularity need different data models, which based on different levels of summarization and corresponding actions taken by the website visitors in the interactive process.[5][7].

5. STRUCTURE PLAN FOR ENTERPRISES’ DATA INTEGRATED BY ODS

There are two structures in analysis of enterprises’ data environment [2]. One is two-layer structure like DB-DW, which is directed at the overall situation’s analysis application. The other one is three-layer structure like DB-ODS-DW, which not only can solve problems of overall situation’s analysis application just like the two-layer structure, but also can solve the problems of daily business demands. We can choose these two constructive strategies of Data Warehouse according to enterprise’s characteristics and specific demands. However, when DB, ODS and DW exist at the same time, ODS has many advantages. For example, we can implement OLTP in enterprise-class and instant OLAP analysis by ODS. What’s more, ODS can simplify data transition interface of DW and its management of data. Although not all constructions of DW needs ODS, the strategies made by it in the three-layer of DB-ODS-DW have the characteristics of integration and enterprise-class. In enterprise management, if information of long-term and daily strategies are both needed, the choice of building the three-layer of DB-ODS-DW maybe better.

Based on data warehouse technology, ODS is a kind of data environment concept which is overall situation consistent and faces to subject. In the process of enterprise information construction, it offers multi-layer data processing environment and builds up the three-layer structure of DB/ODS/DW. In this structure, ODS is the middle layer. On one hand, it includes the overall consisting data and detailed real time data, which can do overall online operational disposal. On the other hand, it is a kind of subject-oriented and integrated data environment, and with small quantity of data. It is helpful to complete data analysis processing of daily decision-making. At present, most enterprises have established perfect database application system, and it is impossible for them to give up all these systems because of the large amount of money in re-investing. So it is more feasible to withdraw data from these application systems to build the ODS, and finally form a technology route of a perfect application system structure. Thus, according to the research of Wei Fang [3], ODS is the best choice for the integration of enterprise information systems and E-commerce data, as shown in Figure 3.

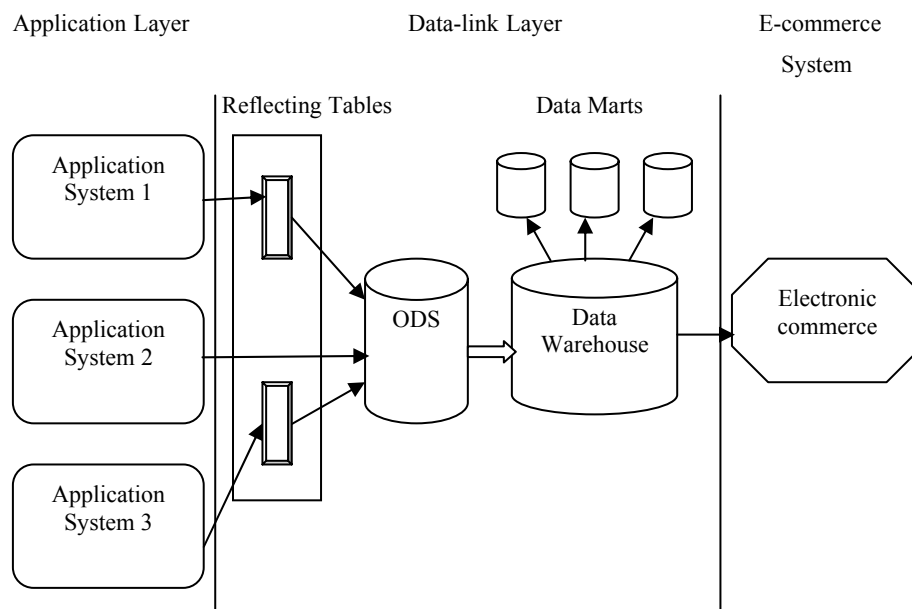


Figure 3. Structure Diagram of ODS and the Application System, Data Warehouse & E-Commerce System

The customer summarization of data and integration can be completed in the enterprise ODS. The summarization and integration of customer data from all enterprise systems and Internet can supply necessary information for successful Customer Relationship Management. In this environment, the data of website information like visiting traffic flow, customer order history (purchasing behavior), customer personal information and customer business activity information etc. can be melted into a data structure, which can offer business group the quick response and

excellent reporting form. Then, individual customer-oriented marketing activities can be initiated, using all these customer information.

In fact, there are two kinds of ODS in the e-commerce environment: Web ODS and Enterprise ODS. They are designed for different purposes [7]. Enterprise ODS is a genuine data processor, which can hold much more data than Web ODS. Web ODS only offers services for internal affairs coming from the web server. Thus, it is a kind of "local" ODS, while Enterprise ODS has real enterprise data which can largely be used for processing. Therefore, Enterprise ODS is a kind of "global" server. And web ODS integrates with Enterprise ODS all day long at different intervals. Giving consideration to the capability and throughput, Web ODS and enterprise ODS are not usually stored in the same server [4].

In the Figure 3, data of the Application System 2 are directly stored in ODS, while data of the other application systems are firstly transformed by mapping table, then renovated, and finally stored into the ODS. The processed data stored in the ODS can be supplied to the data warehouse. These reorganized data which stored in the ODS can be provided to data warehouse. Besides, data in ODS and data warehouse can also be provided to e-commerce systems [3].

6. CONCLUSIONS

Development of modern enterprises needs statistics and analysis, which basis is data integration. With the development of E-commerce and business on the Internet, enterprises urgently need a set of data integration system which is reliable, safe, flexible and easy to extend. Data integration can implement integration and automation of design and production, while lacking of the means to optimize global production process. Strategy-supporting puts emphasis on building up enterprises' optimizing models and strategy supporting systems. And these models can offer supporting for top management to solve problems like sales, price, investment and production plan. With the expanding of enterprise's production scales and fields, also associated with the deep researching on production, the models we build are more and more complicated, and models' dependency relationship to data is stronger and stronger. If these models doesn't base on data integration, accuracy of data and timeliness can't be guaranteed [8]. Data integration plan of enterprise system is based on the rules of practicality and robustness. The suitable plan should be chosen according to the requirements of enterprises system construction. In addition, before choosing the suitable one, we should thoroughly analyze the merits and demerits of the plans. Some people think ODS can partly take the place of data warehouse. In fact, ODS and data warehouse have their respective characteristics, and they have no conflicts with each other, so they can not replace each other. Since ODS is more close to the area of operating systems, it can greatly improve the synthesis of the operating systems. In the process of data integration, if Chinese large enterprises can better utilize the advantages offered by ODS, it will benefit for them and offer them a better environment in their future data warehouse projects.

REFERENCES

1. E.F. Codd, S.B. Codd, and C.T. Salley, Beyond Decision Support, *Computer world*. Number 6, pp.46-50, (1993).
2. X. Su, J. Yang, N. Jiang, and X. Shu, *Data Warehouse and Data Mining* (Tsinghua University Press: Beijing, 2006).
3. W. Fang, C. Zuo and Y. Sun, Realize Data Integration of Big Business, *Computer Engineering and application*. Number 6, pp.218-220, (2002).
4. W.H. Inmon, R.H. Terdeman, J.N. Montanari, and D. Meers, *Data Warehousing for E-Business* (John Wiley & Sons Corp: New York, 2004).
5. Z. Li, *Principle and Application of Strategy Support System*, eds. Y. Liu (Higher Education Press: Beijing, 2005).
6. H. Xia, *Data Warehouse and Data Mining Technology*, eds. X. Chen (Science Press: Beijing, 2005).
7. Z. Li, Effects and Methods of Strategy management analysis technology in MIS Plan Construction, *Science Management and Technology Management Research*. Number 1, pp.211-214, (2005).
8. W.H. Inmon, *The Operational Data Store* (China Machine Press: Beijing, 2001).
9. R. Kimball and R. Merz, *Building the Web-Enabled Data Warehouse* (John Wiley & sons, 2000), pp.47-151.
10. Q. Guo, J. Yu, S. Liu, J. Zhao, and Y. Qi, *Theory and Practice of Enterprise Data Mining* (Huanghe Water Economy Press: Zhenzhou, Henan, 2005)