An Improved ODS Platform on Large Enterprise-level

Chengyu Zheng
China Telecom Sichuan Branch,
Chengdu 610041, China
e-mail: zheng.chengyu@189.cn

Liqun Huang*
School of Software Engineering, Huazhong University
of Science and Technology, Wuhan 430074, China.
e-mail: huanglq2002cn@163.com

Xiaotao Huang*
Network and Computation Center, Huazhong
University of Science and Technology, Wuhan 430074,
China.
e-mail: huangxt@mail.hust.edu.cn

Abstract-With rapid development of big data applications, telecom operators pay more attention to the management and analysis of various data. Presently, as the telecom operators haven't built a uniform enterprise data platform, and application systems and internal data models differ from one another, the complexity of interfaces increases gradually, which restricts the ability of business systems to perform all their business supporting. But now, the ODS increases the maintenance efficiency, decreases IT operations and maintenance costs, improves the quality of supporting data, assures the operational revenues, improves the service quality, provides strong support for operating decision making through using the ETOM model , adopting SOA, applying ODS and EDW technical framework for full service operation and all professional integration according with the general principles, and introducing EAI and management mechanism of message control.

Keywords-operational data store,data warehouse,telecom operator, big data, enterprise data architecture

I   INTRODUCTION

With booming development of the Internet, as of June 2014, China now has more than 630 million Internet users, the Internet penetration is 46.9%, increased by 1.1% compared to that at the end of 2013[1]. According to a survey by Informa Telecoms & Media in 2013, approximate 48% of total 120 operators around the world are implementing big data business which certainly will become a strategic advantage in five years [2]. However, big data and data warehouse are two different aspects of one issue. Big data collects various types of data for comprehensive understanding of customer, while data warehouse which is based on technology manages data during life cycle and manages metadata and monitor data. Big data is an important chance for enterprise to create value, and enterprise data warehouse is inevitable [3].

In recent years, China Telecom has built a EDA framework to manage operating information which consist of MSS, BSS and OSS, etc. to meet changing market demand [4]. Business support systems adopt the independent data distribution policy which means that system contains what kind of data the functions require,

and system data are distributed self-contained. They didn't integrate shared data and make a uniform data view. The data, functions and business of those systems all have many connections. Seeing from the perspective of the end-to-end functional requirements of operators, the implementation of a function usually needs the collaboration of multiple systems. Information synchronization and function collaboration among systems are implemented with interfaces generally. But in practice, the proposal obviously embodies the weakness of passive adaptation to requirements. In usual, an interface is created between two systems according to the business requirements. With the development of business and requirements, the interfaces must be readjusted incessantly. In this way, the interfaces among those systems will be more complicated.

Due to non-uniform enterprised data platform, different system producers, and various inner data model, the complexity of interface is increasing with high-speed changes in business, the whole business adaptability of each business support system cannot exert, and also, it cannot meet the demand of big data. As the overall support system has no specific architecture, applications and data (basically department-level data and systems) are highly coupled, and data can't be shared. Data, protocols and standards are replicated, synchronized or transformed among application systems through interfaces, which results in too much redundant data and repeated functions, low running efficiency, weak system expandability and high construction costs. Thus, sharing data between several systems is an important issue for IT, it is necessary to adopt data share platform among systems. To solve the problems above mentioned and form uniform enterprise's data view, combing with the actual situation of the Telecom enterprise, the enterprise ODS project must be planned so as to build a uniform enterprise data platform to ensure the consistency of data, reliability of data quality and robustness of data models.

II   GENERAL DESIGN IDEAS

ODS[5] is founded to share data so that data from each system can syncretize with each other according to

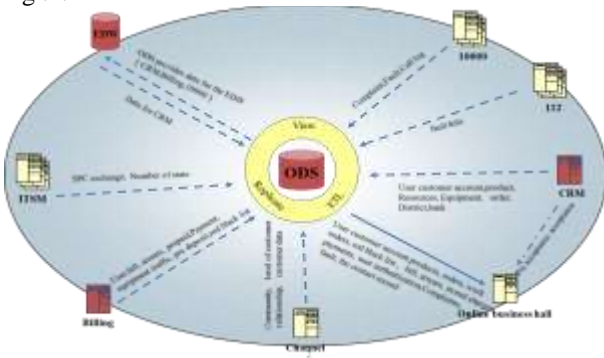data management status of China Telecom, as shown in Fig .1.

Figure 1.   ODS Panoramagram

A new ESA is founded to meet the demand of telecom companies via ODS. Traditional ODS improved the efficiency of data storage of EDW[6], but it cannot make a good balance between magnanimous data storage and operating efficiency of system, as well, frequent input and output caused too much system overhead. However, ODS cannot avoid some disadvantages like flaws exist in data management, meta data in ODS and EDW is not clear, and it lacks dispatching mechanism among concurrent process due to making no distinguish between ODS and EDW. In order to solve those problems mentioned above, ETOM model[7][8] is founded, SOA[9] is adopted, EAI and management mechanism of message control are introduced based on ODS+EDW framework, then uniform enterprise data platform is built, as well as enterprise IT system is rebuilt. So, it is necessary to adjust the existing ODS model, add management mechanism on data storage, namely changing Operational Data Store to Operational Data System. This paper is concentrated on the ODS+EDW+EAI to build a uniform enterprise data platform as shown in Fig .2.
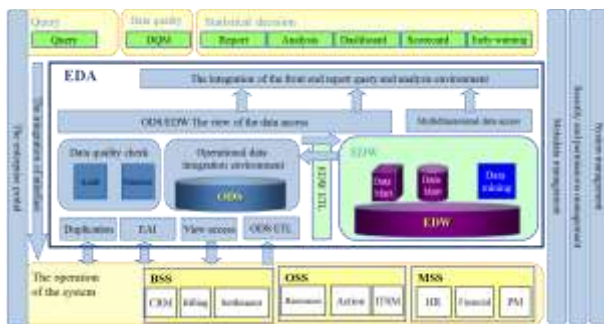

Figure 2.   Concept Map of EDA

Enterprise data framework is consummated through building ODS so that enterprise is able to unify data model and manage operating data and provide uniform support to business. ODS, as data synchronization source and data exchange center, can provide cross – system real-time reports and query statistics application, can implement ODS based data audit and data management, and can build uniform, standardized, subject-oriented, integrated, volatile and detailed operational data platform.

Cleaning and integrating operational data is one major data source of EDW. Add a data integration layer (also named data buffer layer) between historical data store layers of IT system and EDW for data separation. ODS consists of two modes including center ODS and local ODS. Data in center ODS comes from CRM and each local ODS, which guarantee uniform management of ODS. Data in local ODS integrate several operating data and provide application service.

III   ODS ARCHITECTURE

ODS is subject-oriented, integrated and volatile data. It locates between operational environment and analysis environment. As one of important sources of high-quality data offered to enterprise data warehouse, ODS will extract, clean, transform, filter and integrate data generated by application systems (billing system, customer comprehensive management system, settlement system, customer service system, network management system and value-added service platforms), and also provide data sharing service like real-time cross-system operation reports for each application system. In the enterprise's business layer, ODS also provides key support for the collaborative applications which shall utilize both cross-system operational data and related analysis results together.

A.   Architecture of ODS System

The output of the ODS construction is not a concrete product but a ceaselessly enhancing and developing process. The general requirements are as follows: create and partition ODS data fields combing with the actual situation of China Telecom according to the EDM, and considering the urgency of ODS application and the future development demands. The overall architecture is designed as shown in Fig .3, which can satisfy the future development demands of ODS and business, and adapt to the change of enterprise data environment by adding SOA/EAI as a loading and invoking means of ODS data and combing the data loading process of ODS and EDW.

From the modularized architecture based on functions, ODS is divided into three levels and three management supporting parts, including data storage layer, data integration layer, ODS application layer,

metadata management, security management and system management.

Relative independence of the modules: Each level and module has its own features, modes and independence in designing and building.

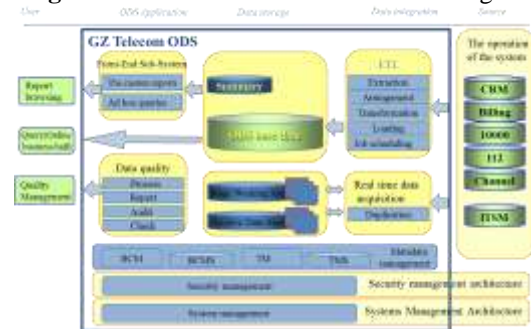**Integration of the architecture**: Ensuring the

Figure 3.   ODS System Architecture

integration of each level and support module from the architecture.

Expandability of the capacity: The mode can be deployed with small configuration and can support enhanced application, management security of ODS in the future.

Expandability of the performance: The levels and modules impact each other and all have correspondingly independent expansion ability. They can be evaluated and expanded respectively according to the characteristics of application in different stages.

1) Data Storage

Considering the data standard and system efficiency, the data storage module is designed according to the hierarchical model and divides into three parts, including operational data storage space offering operational data integration and user query, stage working area and archive data storage space supporting backend data integration, front-end data store space storing generated query result files in the front end.

ODS is the main body of data storage. The main purpose of operational data storage is to integrate enterprise's near-real-time operational data according to the subjects. As the data base to support applications based on ODS, its data model emphasizes the integration and consistency of data, as well as the efficiency of data query.

ODS is a near-real-time data integration environment and offers application functions which require high timeliness and are able to support daily business and management.

2) Data integration

This module is responsible for acquiring data from business system and integrates those data into the data stored in the ODS. In other words, it is a process including data acquisition, data transformation, data loading, job scheduling and exception processing.

In order to ensure that the data which requires being synchronized near-real-time in the ODS can be loaded and updated in time, data replication, EAI and ETL can be used. Considering that China Telecom uses less EAI and has the application experience in the data replication, the data synchronization of ODS will use the combination of data replication and ETL.

3) Application of ODS

ODS are applied to three aspects:

- ODS based business functions like query function of online customer service.
- ODS based data management like data quality check and audit. Data quality management is an important module of ODS system, including an integrated data quality management solution.
- Integrated query and integrated operation report.

Considering the difference of ODS applications, it is allowed to separately develop online customer service and integrated business report according to the features of its own business. The integrated operation reports can implement the integration and personalization of user application interface via the enterprise portal.

4) Metadata management

This part includes business metadata management and technical metadata management, and corresponding metadata rules and procedures as well. Based on the capacity of current metadata software, metadata software emphasizes on data lineage analysis and change impact analysis of technical metadata. But business metadata is represented in the semantic layer of front-end reporting software.

The metadata management is very important for the data application, but also is a relatively weak part in the current technical implementation. Therefore, the metadata management shall focus on management rules and procedures and offer technical support in possible cases.

5) Security management

ODS, as an operational data support platform, shall carry out security management to the user's access rights, data scope available to be accessed and etc. because it integrates operational data from different systems.

Seeing from the current planning of ODS, there are following four cases:

- ODS based business application in other application system: Application system controls customer's access rights.
- Ad hoc query: make right control using database and query tool.
- Integrated report: make right control in the report presentation layer.
- Data exchange support: provide management control for developer and maintenance man.

6) System management

Besides traditional data backup and recovery ability, it is expected that the system management can use system, database and application monitor management mode of the ITIL on the Telecom enterprise to guarantee the continuous running of system.

As shown in Fig .3, traditional ODS was built to improve the efficiency of data storage of EDW[6], but it cannot make a good balance between magnanimous data storage and operating efficiency of system, as well as, frequent input and output caused too much system overhead. However, ODS cannot avoid some advantages like flaws exist in data management, meta data in ODS and EDW is not clear, and it lacks dispatching mechanism among concurrent process due to making no distinguish between ODS and EDW. In order to solve this problems mentioned above, ETOM model[7][8] is founded, SOA[9] is adopted, EAI and management mechanism of message control are introduced based on ODS+EDW framework, then uniform enterprise data platform is built, as well as enterprise IT system is rebuilt. So, it is necessary to adjust the existing ODS model, add management mechanism on data storage, namely changing Operational Data Store to Operational Data System.

C. ODS Data Model

In existing ODS projects, ODS system get data from integrated business system, 10000 system, integrated settlement between networks, real-time billing system, and channel support system. It provides data support for online business halls and EDW system. Data interaction information will be referenced to define data scope when data model is designed.

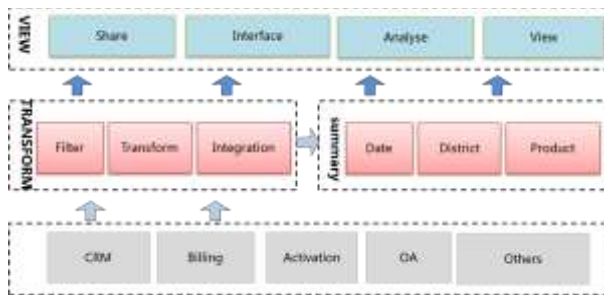The data model is divided into 4 levels as shown in Fig .4:

Figure 4.   ODS Data Model

1)   Source data layer

This layer is same as data architecture of source system to keep the integrity and consistency of data and source data. Its main features are as follows:

- The data stored in the source layer are data from application systems, including billing system, customer comprehensive management system, settlement system, customer service system, network management system and value-added service platforms.
- Data model in the source layer is same as that of application systems.
- Data fields in the source layer synchronize real-time or near-real-time with data of application systems.

2)   Summary data layer

This layer stores statistics data and summaries data formed according to the subject dimension, which are coarse-grained data formed after summarizing data by business and performance in the source layer. Its main features are as follows:

- Form summary data according to the subject report processing requirements.
- Store summary data according to the subject. Summarize operational data according to such dimensions as date, district, product specification, commodity specification, channel type and processing type, and store such summary data.
- Continue to summarize on basis of old summary data to form multi-level summary data.

3)   Transform data layer

This layer cleans and transforms data from source layer according to the format of ODS data model to form standard data format. It is the core data layer of ODS. Its main features are as follows:

- Transform data layer is the core data layer of ODS.
- In principle, the data in the transform data layer are those with uniform coding format. They can be considered as enterprise data standard used

to guide peripheral systems to unifying data format gradually.

- The storage data model conforms to the CTG EDM. The data are organized according to the subject of data.

4)   View data layer

This layer stores data integrated in ODS and offer data sharing. In principle, the data storage is not offered. Its main features are as follows:

- View data layer offers external systems shared data.
- The data shared to systems shall be mapped and used in the view data layer if possible.

IV   SUMMARY

A new ODS technical solution is founded to meet telecom enterprises' demand,. This solution builds uniform enterprise data platform through introducing EAI and management mechanism of message control, using the ETOM model , and adopting SOA, with the general principles on the base of traditional technical framework of ODS and EDW. ODS is divided into three levels and three management supporting parts, and system functions are hierarchized and modularized meeting the demand of unifying data by telecom enterprise. This new solution has been operated more than ten years, and achieves the desired aims. The ODS increases the maintenance efficiency, decreases IT operations and maintenance costs, improves the quality of supporting data, assures the operational revenues, improves the service quality, provides strong support for operating decision making.

REFERENCES

[1]. Top ten trends of Chinese Internet Development in 2015

[2]. http://www.leiphone.com/news/201412/mX6cHZXIjTJIiapJ.ht ml 2014

[3]. Zhihua Fu, the application of big data in the telecom field,

[4]. http://www.leiphone.com/news/201411/eAYu5wZWfAKAYvr G.html 2014-11-17

[5]. Albert-László Barabási, Bursts[M], Beijing, China Renmin Universality Press, 2012

[6]. Urban Planning Informatization, 2013 No.3 preface, http://www.wpl.gov.cn/pc-258-49724.html 2013

[7]. Zipei Xu, Big data: data revolution to be coming [M], Guangxi Normal University Press, 2012-07-01

[8]. Big data and enterprise data warehouse http://labs.chinamobile.com/mblog/1052295_204838, 2013-8-6

[9]. Author lnmon W H, Translator Zhihai Wang etc., Data Warehouse (4th Edition), Beijing, China Machine Press, 2006

[10]. China Telecom, CTG-MBOSS EDA Specification V1.0, 2005

[11]. Xuemei Hu, The application vision of Data warehouse in China Telecom, http://www.mscbsc.com/viewnews-45047.html, 2010

[12]. Yun Ma, Alibaba is not an Internet Company; the target of Taobao is to acquire data, http://www.kejixun.com/article/201412/84969.html, 2014-12-1

[13]. Hequan Wu, Big Data Thinking, Science and Society, 2014 Vol. 4 NO.1, pp 1-13

[14]. Xinming Chen, Construction of data warehouse of Taobao data platform [D], Dalian University of Technology, 2013.4-10