Research on Airport Data Warehouse Architecture

WANG Jian-bo

FAN Chong-jun

Business School
University of Shanghai for Science and Technology
Shanghai 200093, P. R. China.

Abstract

Domestic airports are accelerating the construction of business intelligence systems, and data warehouse is core of the airport decision-making system. On the background of Shanghai Airport，This paper analyses the key demands of airport BI and data warehouse, discusses the application architecture of airport data warehouse based on the comparison of several other types of architecture and explores the function architecture of airport data warehouse architecture. This paper aims to provide the reference and help for the research and planning of relevant airport projects.

Keywords: Data Warehouse architecture; airport; data mining

1. Introduction

The concept of data warehousing dates back to the late 1980s. Several different definitions have been used to describe data warehouse. McFadden and Watson[1] define a data warehouse to be a collection of integrated, subject-oriented databases, designed to support decision enhancing activities, where each unit of data is relevant to some moment in time. DW (Data Warehouse) is used to collect data designed to support management decision making.

So far, the foreign scholars have done a very in-depth study on data warehouse. McFadden and Watson[1] present one of the earliest architectures for data warehouse design. At home, Zhang Shuming[2] discusses the basic principles and main design methods of data architecture through a data warehouse system implementation case of a telecommunication enterprise, yet research on Airport areas are still less.

The data warehouse as an important aspect of the construction of the airport business intelligence systems is an important means to effectively support airport decision-making, data warehouse architecture design as a fundamental basis in the data warehouse technology is a description of the system architecture and modeling directly determine the applicability and validity of the data warehouse system, an extremely important role in the process of building data warehouse system. This article rely on the data warehouse system construction projects in Shanghai airport, have a brief discussion on data warehouse architecture such as Bus architecture, Hub-and-Spoke architecture, Federated data architecture, and proposed a data warehouse architecture for airport development program.

2. Data Warehouse

The purpose of the data warehouse is to build integrated data analysis-oriented environment for enterprises to provide decision support. In fact, the data warehouse itself is not "production" of any data, at the same time itself does not need to "consume" any data, data from an external source, and open to external applications. The basic structure of the data warehouse contains data inflow and outflow of the process, which can be divided into four layers - the source data, data staging area, data presentation area, data access tools. Operational Source systems: Data (Structured and unstructured data) from OLTP systems, database such as SQL, Oracle, Sybase and other modern data source XML, WSDL, LDAP, JDBC etc.; Data staging area: a DW environment includes an extraction, transformation, and loading (ETL) process; Data presentation area includes the data warehouse and an online analytical processing (OLAP) engine; Data access tools: including some Query Services like Warehouse Browsing, Access and Security, Data Mining, Reporting, Activity Monitor Data Mining etc.. DW logical architecture shows as Fig.1.

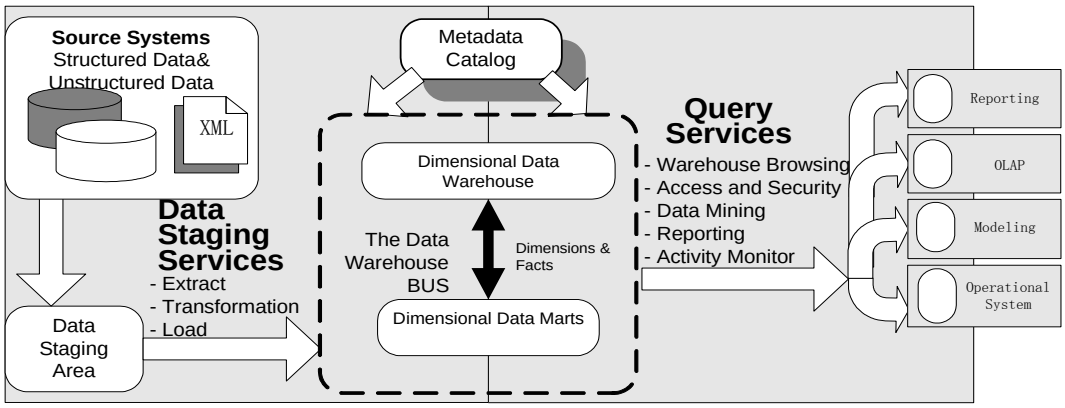

Fig. 1 Logical architecture

3. Data Warehouse Architecture

3.1 Data Mart Bus Architecture

Bus Architecture[3] starts with a business requirements analysis for a specific process such as orders, deliveries, customer calls, or billing. One mart is created for a single business process, Additional marts are developed using the conformed dimensions of the first mart. The data warehouse is a joint constitute the data mart, data passed via the data bus.

The data mart established according to the appropriate standard for data warehouse that same dimension. If there is no adhesion of the consistency dimension, data marts can only be an isolated application. Precisely because of the previous bus architecture and consistency dimension, this architecture can guarantee to ensure consistency of business data but also in the process of the gradual establishment of a data mart. In this method, you can understand the data mart set for the logical child of the entire data warehouse system. The data warehouse bus architecture diagram show as Fig. 2.

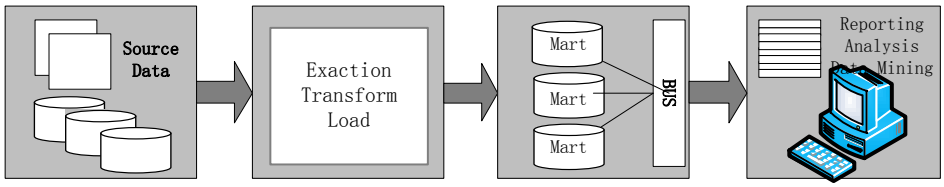

Fig. 2 Data Mart Bus Architecture

3.2 Hub-and-Spoke

The Hub-and-Spoke data warehouse architecture is to collect data from various source systems and data available to the various data marts and mining warehouse. This architecture develops after an enterprise-level analysis of data requirements, mainly focuses on building a scalable and maintainable infrastructure and consists of a centralized hub that accepts requests from multiple applications that are connected through spokes. See (Fig. 3). Integrated data distributed to each data mart from the central data warehouse.

This architectural approach is of course, disadvantages, although data marts build according the integrated central data warehouse, yet they can not communicate between each other. At the same time, due to the establishment of the data mart, ETL program is different, so the data among the various data marts may be inconsistent, which makes follow-up construction of this architectural approach becomes complicated.

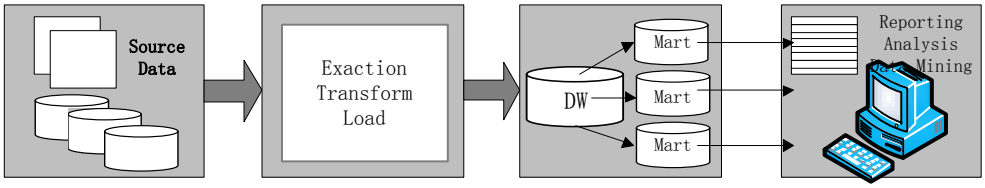

Fig. 3 Hub-and-Spoke Architecture

3.3 Federated Data Warehouse Architecture

The federal data warehouse leaves existing decision-support structures and data marts in place, and shares information among a number of different systems. The data is either logically or physically integrated. The federal data warehouse diagram show as Fig. 4.

Federated data warehouse architecture regardless of geographical federal or functional need to build data warehouse respectively, and this can easily lead to integration is not complete unless the federated data warehouse architecture is also used similar functionality as the Bus Architecture, and retain conformed table in the data preparation and continually update it. Of course, for a quick solution to the original independent data mart data exchange federal architecture has some advantages.

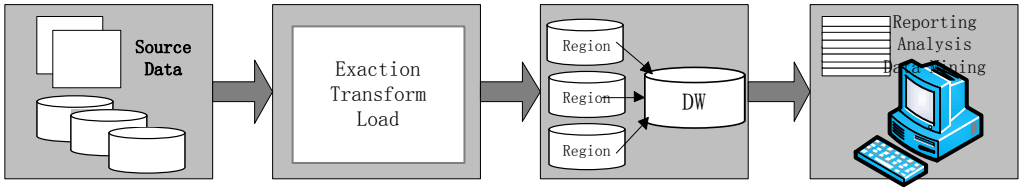

Fig. 4 Federated Data Warehouse Architecture

4. Airport Data Warehouse architecture

4.1 Airport information systems and demand analysis

After years of information technology process, the types of airport information systems, covering a wide range, overall, include the business systems and management information systems. The production system is divided into the flight information system, security monitoring systems, equipment and facilities management related systems. Management system, including purchasing inventory management system, property rental management system, asset management system, contract management, tender management system, human resources management system, investment project management system, customer relationship management systems, financial management systems, logistic Management System, etc.. Due to different period of development of information, there are some of the same information system and part of the independent system between the two. As businesses grow, large amounts of data exchange happen between the two, as well as The Group and Air Traffic Control, and some other external information systems. Between the current data exchange is a peer-to-peer mode (Fig. 5).

Shanghai Airport Group business intelligence services to senior management and decision making, supplemented by functional business support. Shanghai Airport, including two airports, Pudong and Hongqiao. The Group does not directly make a specific production business, but mainly responsible for the possession of business entities under control. To this end, data warehouse architecture should be beneficial to senior management, financial management, strategic management and statistical analysis, performance management problems.

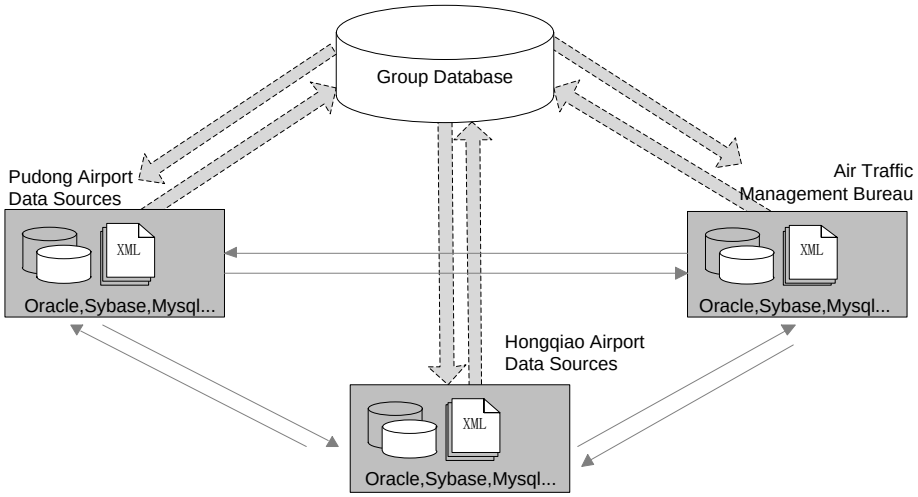

Fig. 5 Airport data exchange mode

3.2 Comparison of the data warehouse architecture

The bus data warehouse architecture is to establish a data warehouse from the bottom, which established in accordance with the business process data marts, and achieved enterprise-level consistency through the data bus and consistency dimension. Its architecture is simple, at the same time, using the dimensional modeling approach has some advantages compared to ER modeling. The bus structure has less effects on the source system data, and some other advantages such as low cost, fast deployment etc..

The Hub-and-Spoke architecture establishes a data warehouse from the top down. According to the established enterprise-level data warehouse, gradually establish department-level applications. Relative to the bus data warehouse, the architecture process is more complex and higher data integration, to some extent, this architecture can affect to the business system data. In the project implementation stage, the hub architecture needs high start-up costs, but the subsequent development of lower cost.

The proposing of the Federated Data Warehouse Architecture mean to solve the problem of information silos based on the use of existing facilities. When we need to build a data warehouse, some companies may already have some of the data mart. However, this data warehouse is often a remedy, so its data integration, data warehouse construction is not thorough enough.

Many factors potentially affect the architecture selection decision, such as upper management's information needs, strategic view of the data warehouse prior to implementation, compatibility with existing systems, perceived ability of the in-house IT staff, end user requirements, emerging technologies. On the basis of analysis of the above factors, combined with the characteristics of the various data warehouse architecture, we believe that the bus data warehouse architecture is suitable for the requirements and future development of Shanghai Airport.

3.3 Airport Data Mart Bus Architecture

In this structure, The DW system extract data from business systems and establish financial analysis mart, statistical analysis mart, marketing analysis mart, KPI management mart, data mining mart etc. (Fig.6). The airport will be based on the phased establishment of specific business area data marts, and each phase of the establishment of a data mart can be mutually compatible and can be combined to achieve a data warehouse.

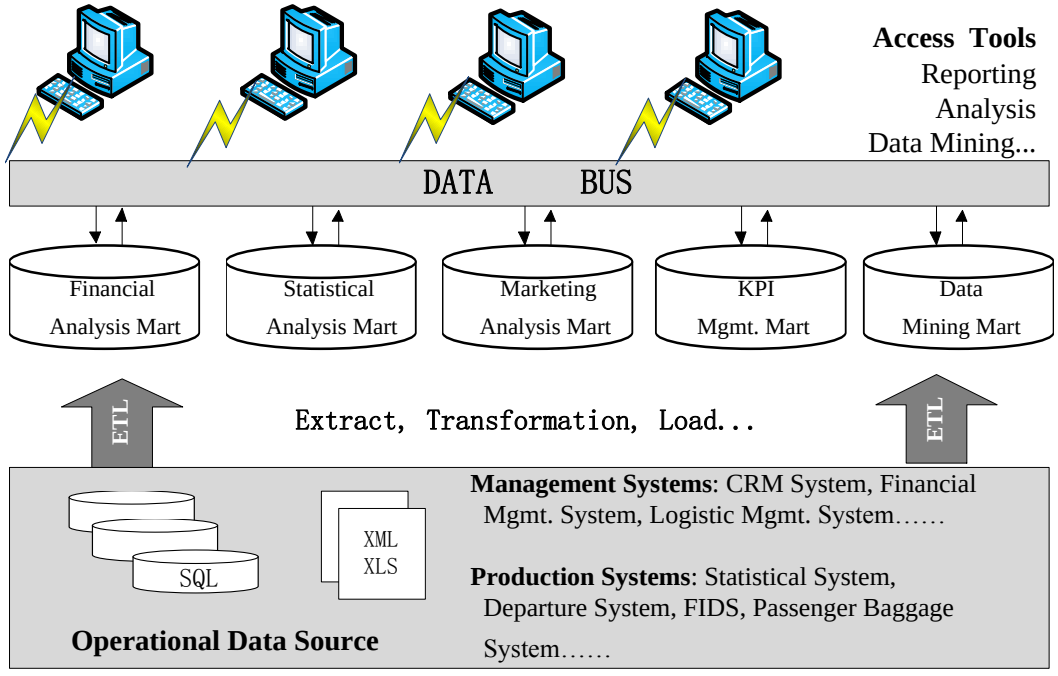

Fig. 6 Airport Data Mart Bus Architecture

The data bus mode is also conducive to the airport in accordance with the development of the business needs, instead of a one-time construction, which provides the conditions for the data warehouse scalability, and also space for business development.

These data warehouse construction methods meet the requirements of the strategic development of the Shanghai Airport, also in line with the step-by-step implementation of the strategic considerations of the development of information technology, and a valuable solution.

5. Conclusion

In this paper, we conduct a comparative study of three common data warehouse architecture according to the needs of decision-making theme of the Shanghai Airport Group. We believe that the data bus relative to the other two type of architecture has a distinct advantage. The data bus architecture model allows businesses to ensure consistency of data dimensions, and the gradual completion of the construction of a data warehouse in accordance with the business development needs. The solution proposed in this paper can provide a reference for the airport data warehouse construction, but also has a certain degree of universality can learn from other industries data warehouse construction.

References

F McFadden, H.J Watson (1996), The world of data warehousing: issues and opportunities Journal of Data Warehousing, 1(1):61–71.

Zhang Shuming (2009), Discussion of Data Architecture Design for Data Warehousing: Information and Communications Technologies, 6:11-15.

Thilini Ariyachandra, Hugh J. Watson (2006). Which Data Warehouse Architecture Is Most Successful: Business Intelligence Journal, 11(1)4—6.