Service-oriented Enterprise Management Information System based on Data Warehouse Optimization and Distributed Computing Method

Yuhong Qian

School of computer technology and software engineering, Wuhan Polytechnic, Wuhan, 430074, China

*Corresponding author(E-mail: yhqian99@163.com)

Abstract

In this paper, we theoretically analyze and review the current service-oriented enterprise management information system based on the large-scale warehouse based on optimization and distributed computing methodology. There are many ways to management information system development from the current trend with structured and modular development method can get the best development effect. Structured design can be done at the global level first system of overall planning and correlation analysis between each subsystem, determine the clear development goal, step by step in accordance with the planning content, and then analyze the system logic structure, ensure that the integrity of the function of each subsystem and level. Modular development can guarantee the compatibility of information transmission and sharing between modules and unity, enhance the flexibility of system development to avoid a certain functional problems affect the system as a whole. Traditional data warehouse system, through to the historical data aggregation and analysis, can provide enterprises with such as next year will adopt a marketing strategy and strategic decision support. And the continuous development of customer demand, also poses a challenge to the technology, the real-time data warehouse in traditional data warehouse on the new data warehouse architecture has been widely recognized. Our research adopts the distributed structure and the data warehouse storage pattern to help enhance the effectiveness of the system and modify the data security and flow control steps to achieve better performance.

Keywords: Service-Oriented Enterprise; Management Information System; Data Warehouse; Distributed Computing; System Modification.

1. Introduction

Management information system development is relatively complicated, it includes the theory of the knowledge management, system, organization structure, the contents of many aspects, such as computer technology, the construction of management information system is a big topic in the field of application in each big enterprise has extensive adaptability. User scope is quite wide. That must be followed for the development of management information system is rapid, easy to modify and easy to maintain, easy to expand, such as the principle, to fundamentally guarantee the speed and quality of software development, first of all need to adapt themselves to the choice of the mode of development. In the reduction of information system input at the same time, guarantee the quality of the management information system, shorten the system development cycle which will improve the effect of system management. The following the first introduced the management information system development pattern, and then analyzes the different development mode of main features. Object oriented from the stable factors an entity of the objective world, system structure, system is described by using the object of the entity. Packages will entity attributes and related operations in the object, class is used to describe a set of objects with the same characteristics, represent the whole class object. The object and class formed the basis of object-oriented component and it is the general basic mechanism method, messages, and inheritance, improve the level of abstraction of software development and improve the software reusability is the basic means for object-oriented program design, focus on the classes and the general class hierarchy on the design, implementation, and reuse of is object-oriented program design and the essential difference between the traditional basic programming. By using the object-oriented method, therefore, the analysis of management information system can make the basic management information system has a good scalability, reconfigurable and compatibility in sum.

In summary, traditional development pattern could be separated into following parts. (1) Network technology is increasingly mature and the concept of a LAN technology was applied to the management information system development and design and then derived the W/S model. It to the whole system in different links, required for different business exchanges and sharing of data resources in the form of files stored in the system node server, all branch work site in the system can get data through the node server resources, compare the stand-alone mode, has its significant share the advantages. In W/S mode, the system of data is not exist in isolation, it has the characteristics of communication and sharing, fundamentally reduces the repetition rate of the data, simplify the system, but also there are some problems. W/S model the overall structure is simple, but

poor expansibility. (2) In the C/S mode, the information system not only has its data processing function, also have the function of file storage, server can be implemented with different division of labor and cooperation, work site is geared to the needs of customer service mode of the port. It will management information system is divided into two parts, namely the client and the server, the server for data management process to provide public services and support, while the client is mainly to deal with local storage data resources, and through interaction of the two servers, information management and the sharing. Its response speed is much faster, mainly through service machine directly communication with client, no transmission intermediary, at the same time has its distinctive personalized features, simple and convenient. (3) In B/S model, the client/server without any business data stored information, don't need to synchronize data, all management information stored in the database server, compare the stand-alone mode, the security is higher. With a standard network topology protocol, can good interoperability and combining with enterprise network view, directly into the IE browser has the better extensibility. It was established based on the network, customers also have permission to access the management information system of network resources, network users can access the system through the network at the same time within the data, information resource is rich, and sharing degree is high, operating at the same time also is relatively simple. But there are some shortcomings, mainly because of strong commonality, certainly will cause its individuation degree is low, mainly in the mouse as a mode of operation which can better meet the needs of rapid operation, page data amount is larger and the program corresponding speed is slow.

Management information system are widely used in the general management of the affairs, strengthened the management of the affairs, make the management work has improved in a wide range of industries. Management information system is a service nature of the system, management information system establishment of the new technology is to create a new type of management mode, make the management model to keep pace with the times, follow the footsteps of the development of Chinese modernization. Modern break the traditional mode of management information system, and greatly improves the efficiency of management, from all walks of life to modern standardized management make opened up a highly efficient mode of management information system. Security problem of computer management information system is the core issues for the systems. The following parts denote the problems. (1) The hacker attacks. The computer management information system of storing a large number of sensitive documents and data which is easy to become target of ulterior motives, especially government agencies office network, and even international hacker attacks. (2) The insider improper use or vandalism. Due to inadequate training or staff for safety value degree is not enough, tend to cause the staff in the process of using improper operation, not only cannot give full play to the utility system, also can cause damage to the system, even sensitive data, so as to pose a threat to system security. (3) The user authentication. User authentication is the first barrier of the whole system security, how to effectively control the illegal user login and use is very important to safety. If the improper control, may cause serious information leak problem. (4) Legitimate user rights allocation. The computer management information system should be to legitimate users according to the different status, positions, post set up different security levels, make its have different permissions. In the figure one, we show the core challenging components of management information system.
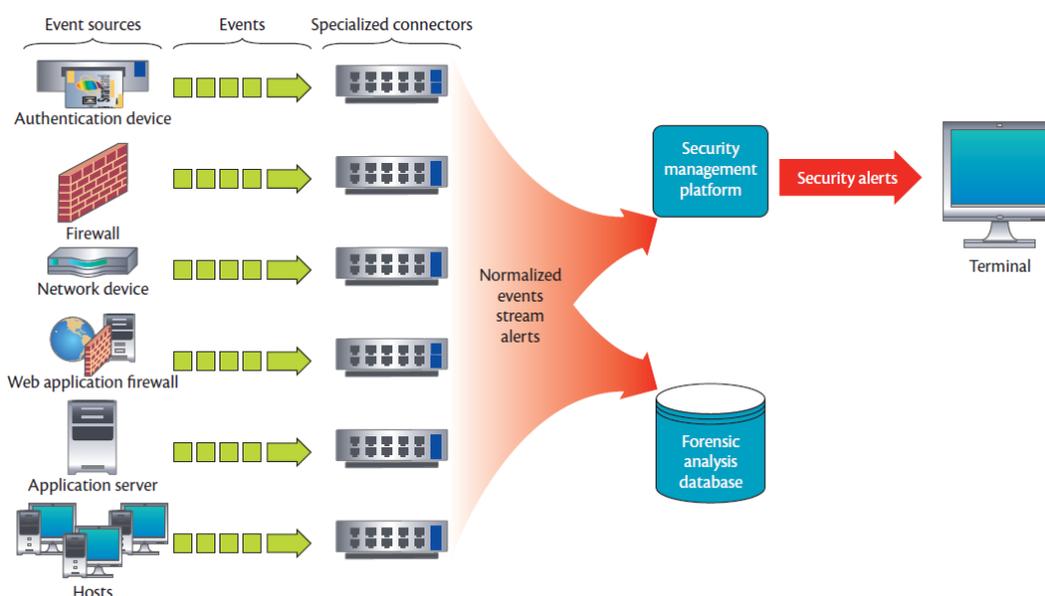

Figure 1. The Core Challenging Components of Management Information System

To optimize the current management information system, we analyze the service-oriented enterprise MIS based on warehouse optimization and distributed computing methodology. After many years of practice of the computer application, the enterprise has accumulated a lot of original production data and all kinds of business data, these data truthfully reflect the enterprise main body and various business dynamic business environment. But due to a lack of centralized storage and management means, to how to make full use of these data effectively, but there has been no good way to solve. Therefore, to make full use of the accumulation of the data resources to provide decision support for enterprise operators and decision makers which is a very urgent and challenging task. The data warehouse from data organization to support the analysis of processing and application oriented database system has bigger difference. It decided the design method of the data warehouse system different from traditional database system design and development methods. For the application of the distributed computing environment, synchronous communication is an indispensable and important way of communication between tasks and homogeneous cluster parallel machine application on computer, usually use the interaction between asynchronous transmission way realization process. Distributed computing technology is one of the important products in the development of the information age today and it allows computer equipment at the same time open a number of services and realizes computer data communication through the network. Distributed computing is the main research field is roughly divided into distributed computing environment and distributed operating system two aspects of study. In the following sections, we will build up our designed MIS based on the prior technologies.

2. The Data Warehouse Access Optimization

2.1. The Introduction of Data Warehouse

Underlying storage structure most used traditional data warehouse storage, will each record is stored in the media as a whole order. Due to the differences between the different data sets of data model, it is difficult to find two identical records and so on the repetition rate of the original data storage level is not high, not easy to reuse. At present, the line store often considers the reuse of intermediate results generated based on a view or query. By reusing connection, projection, grouping, such as complex processing of the intermediate results and meets the timely response to the user query and the related requirements of on-line analytical processing and other applications. Although this strategy can improve the query efficiency to some extent, but due to the application contains huge amounts of data, analytical data query execution process of the intermediate results can be repeated access also has characteristics of "mass", in the case of the limited buffer, easily happened memory contention problem. The design of the data warehouse generally go through needs analysis, concept design, logical design and physical model design four stages. Among them, the core conceptual design is to determine the correct implementation of the key, request the conceptual model specification and complete into a logical model won't appear when the fuzziness and ambiguity.
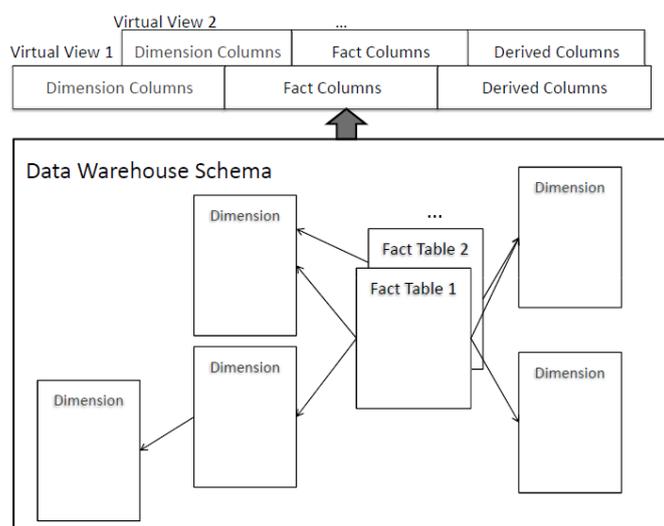

Figure 2. The Structure and Organization of the Traditional Data Warehouse Schema

In the above figure two, we demonstrate the primary structure and the organization of the traditional data warehouse schema. The basic data warehouse organization and step is shown as the follows. (1) Extraction of data extraction is mainly for distributed database, after more clearly understand the meaning of data, planning the required data sources and the operational data sources to read the rules and principles of the incremental extraction, and to the specified destination export work. Source data is usually stored in a distributed way, and

the type of diversification, may be a relationship between plough databases which could be a text file. (2) The data transformation is the most complicated part of process. The reason is that the data transformation to a variety of data makes a unified standard, the standardization is put in storage of data, therefore, involve methods and techniques are more. It is mainly aimed at establishing data warehouse model, through a series of transformation to realize the data from the business model to analysis model, on the basis of the built-in library function with the expansion of the custom scripts or other way, implement various complex transformation and support for debugging environment, clear monitoring the status of data transfer. (3) Load the data warehousing. The main task is to load data the steps above and the high quality of the processed data into the data warehouse, data warehousing. The way we can be loaded directly by the data file or direct connect to the database for data loading.

2.2. The Data Warehouse Access Optimization Algorithm

The scattered, chaotic, history and the current data set up to establish the data warehouse, excavated from the established data warehouse this part of the create benefits for the bank customers, build a model from the complex customer information, records the information to the customer for dynamic tracking and monitoring. The calculating customer value, locking specific customers, analysis of the potential customer base, product configuration according to the customer the value of the selected services, creating profits and customers to establish long-term relationship is the essence of optimization. To finalize the optimization procedure, we analyze the issues from the following perspectives. (1) The consistency of query results. Due to the persistent data arrival and dynamic query, can appear the same query requests by the influence of different query condition and different time inconsistency in the query results which need to query the consistency of process management. (2) The data of the novelty. The real-time data must be loaded into the system in a timely manner, to support the tactical query analysis time consistency in the continuous data integration environment, there will be a time-series data inconsistency phenomenon, time consistency management should be carried out establish time-series model. (3) Take the initiative to the timeliness of decision-making. Real-time system can capture a variety of decision rules limit movement and make response.

Through the above analysis, we found that the real-time data warehouse has a very broad market demand, in line with the enterprise information process, worthy of study and application. Although the real-time data system has some disadvantages and need to combine some of the advantages of traditional data warehouse is improved. To research the change of rapid data capture method, realize the source data changes in the system for selective positioning and capture, satisfy the requirement of zero latency, minimize the invasion of the source system, reduce the load of the source system, to ensure source system performance is not falling. To focus on continuous and efficient data loading technique to realize real-time and efficient continuous data loading. Implementation of the data cleaning and transformation process contains internal link to carry on the reasonable and the effective organization, to improve the rate at which data processing and concurrent degrees.

At the same time according to the different requirements of users for data quality, data of load, reasonable allocation of system resources, improve the data load performance. In the figure three, we show the solutions of the data warehouse optimization method. To focus on the efficient data distribution mechanism, make each captured data changes after the message queue, completed by message queue data distribution, ensure consistency and integrity of the message transmission and at the same time effectively maintain data dependence and time dependence of transactions. Data warehouse to support a variety of possible query, but there are a lot of queries is very complex, involving a large amount of data, and to select these data, projection, connection and other complex processing. This process in the traditional database query time overhead is very large and it should not be tolerated in decision support system. Therefore, the data warehouse for possible queries to connect to the original data, projection, grouping, such as pretreatment, and will be in the middle of the data saved to the physical disk, when faced with the same query which can be directly reused materialized in the results.
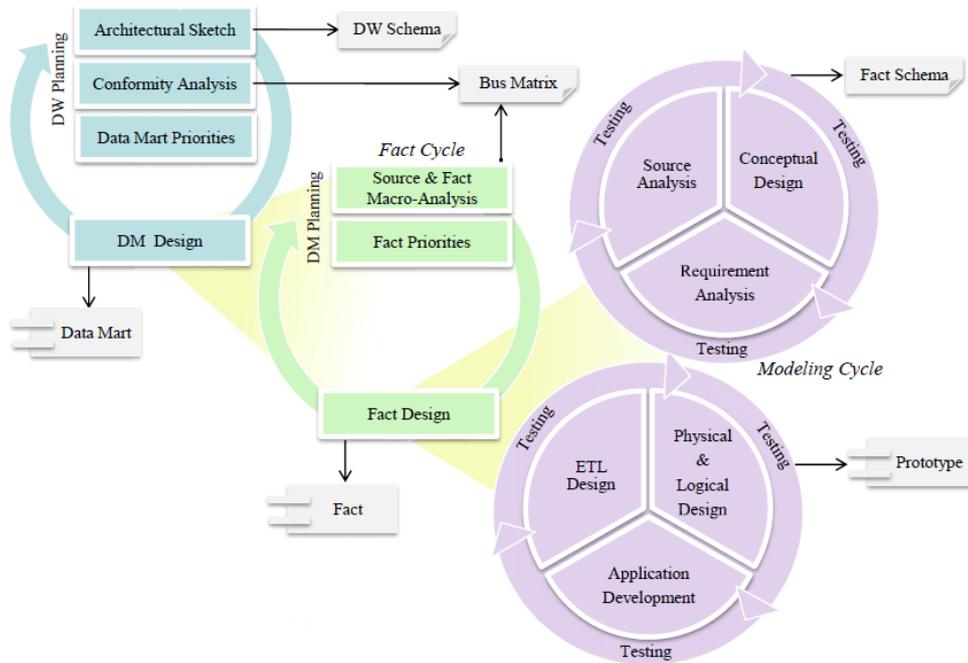
Figure 3. The Solutions of the Data Warehouse Optimization

3. The Distributed Computing for MIS Data Warehouse

3.1. The Principles of the Distributed Computing

In a distributed computing environment, a variety of heterogeneous resources in different management domains are subject to the different management strategy, so it is difficult to undertake collaborative allocation for them. In order to guarantee the task from start to finish can obtain the necessary resources, can adopt the method of reserve resources. So-called resource reserve is that before the actual use of resources for a period of time of booking in the future, to ensure that users in use process to obtain the required resources. Distributed computing technology is one of the developments of computer science and technology, the main working principle is through multiple computers distributed connection of comprehensive treatment of the data, through multiple computers strong work ability to break down complex problems, solve some calculation problems. One of the most core content is to be able to calculate program to find the most appropriate computer to complete the work. At present, the computer in the basic field of distributed computing technology with more than hundreds of, but most did not closely linked and the lack of system management and industry regulation technology is not good for the future development. In the following formula one, we define the general structure of the distributed computing system.

$$D = (P, Ch), P = \{P_1, P_2, \cdots, P_n\}, Ch = \{C_{i,j} : \forall 1 \leq i, j \leq n\} \tag{1}$$

The corresponding time complexity could be summarized as the following equation.

$$\sum_{d=d\min}^{d\max} \sum_{t=tstart}^{tstop-d} d = O(n^3) \tag{2}$$

For the leave request in real life, have a plenty of can expand, and others cannot expand. If a resource manager is extensible received request, you can use the above method can expand the reserved bandwidth or duration are adjusted to meet the request. However, if a resource manager cannot expand request received and don't have enough resources, application time only rejected the request. The corresponding projection could be summarized as the follows.

$$f : \{0,1\}^b \times \{0,1\}^b \rightarrow \mathbb{R} \tag{3}$$

Time slot each array element in the array represents a time slot, the value of the element represents the allocated resource in the time slot. Time slot array to choose a suitable length of the array, the value reserve to ensure that the request of the end time can't beyond the boundaries of time slot array, try to save the basic storage empty asked again at the same time which need to be asked to find a balance point of the two.

3.2. The Distributed Computing based Management Information System Data

Parallel computing is the foundation of cloud computing. In general, use a variety of computing resources at the same time, execute multiple instructions, the collaborative process to solve the problem of big calculation, the calculation model is called parallel computing. Parallel computing can be divided into parallel and spatial

parallel asked. Due to the time when parallel refers to the executable program multiple instructions overlap, has the nature of assembly line work, therefore, precise, it is a quasi-parallel implementation technology. With the reliability of the network has the following calculation formula of the reliable nodes.

$$R(G) = p_e R(G \times e) + q_e R(G - e) \tag{4}$$

Overlapping messages can arrive on time, and each receives a message, receiving returns a response message to the sender, as compared with the data of the reply message size is very small, so the reply message transmission time ignored which is defined below.

$$r(v) = \max \left\{ \min_{p \in \text{Ped}} \{ s(p_i) + \beta(p_i, v) \}, S(u) \right\} \tag{5}$$

$$R(v) = \max \left\{ \max_{p \in \text{Ped}} \{ s(p_i) + \beta(p_i, v) \}, r(u) \right\} \tag{6}$$

From the basic work of distributed computing model, a central server will be a large computing tasks divided into several or more task unit, using the Internet assigned to multiple computers to complete respectively, after the completion of the task and returns to the central server. This kind of work mode determines the characteristics of the distributed computing. First, the system's overall coordination, at the same time, also allows the independent computer temporary problems or failures and distributed computing tasks of each computer are only a partial perspective, for the whole system, its influence limited. Ask the related node communication is given priority to with message passing, data transmission volume is small, also can't see the global, local nodes of each node has its own processor and main memory.

System structure is a dynamic variable, it depends on the quantity, scale, structure, computer network delay as the task and system operation, type of computer and the network connection may change, may even for executing a distributed program that alters system. As the network can span range is very wide, if the scientific and reasonable design, distributed computing scalability is necessary. Our proposed optimization could be summarized as the follows. (1) The resource distribution in the number of sites scattered on the physical and on the basis of transparent to users to realize resources sharing to make the single user resources available exponentially. (2) Scattered resources unit can collaborate in the distributed system, to solve the same problem, under distributed operating system control, achieve the function according to duplicate or task resources according to time overlap of different forms, such as parallelism. (3) The redundancy in the system resources and autonomous control method make the system with dynamic reconfiguration ability, even if the system is local failure can also continue to work. So has the reliability and fault tolerance. (4) Module as the system expansion or resource updates the increase of the unit, not like a centralized system to replace a large part of the whole system or change system and the system configuration is easy to change. (5) Distributed computing resource units form the relatively independent modules and they are the Internet connection into a single system. Modules within a certain range of increase or decrease or replacement will not affect the integrity of the system. In the figure 4, we illustrate the distributed computing based optimization for the system.
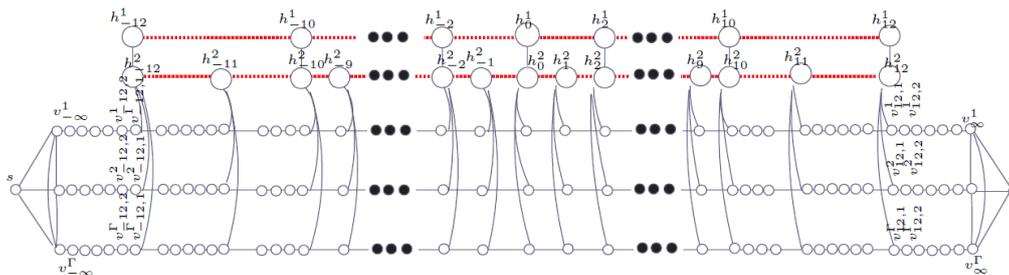

Figure 4. The Distributed Computing based MIS System Structure Modification

4. The Service-Oriented Enterprise Management Information System

4.1. The Traditional Management Information System

Information system to the influence of the internal structure of enterprise organization in the process of development, in order to realize its strategic target and development direction, must constantly according to the functions and powers and responsibilities of the internal personnel division of labor, arrangement, streamlining and so on a series of adjustment, this relationship is formed by the enterprise's organizational structure. With the establishment of the concept of information resources, its development and utilization become a strategic task of enterprises, more and more enterprises have set up information management agencies, large scale and the importance is also growing. In order to ensure the effective management of information resources and the full development and utilization, but also opened a job and work. This trend more evidence and established a place in the enterprise information system. With the development of the network, the network information system also

uses the advanced technology for information exchange and information sharing. Information is not restricted by geographical and time in such a favorable conditions and environment.

Enterprises in order to adapt to the market demand rapidly changing and increasingly competitive environment, mobility requirements for the production and business operation management activities can be done by flexible organization structure. Refers to the enterprise management information system with the aid of modern information technology, optimize the enterprise management process prompting companies more efficient. Previous to enterprise management information system research mainly focuses on the combined with the modern enterprise, improve enterprise efficiency and the system itself design here from management information system and enterprise memory, learning and discuss the perspective of enterprise pressure. With the development of computer industry, the modern management information system on the application of the computer technology has become more and more broad prospects, network and database technology will have a very good improve, along with the progress of computer science and technology in some technology also has some big span, this means that the management information system can be used in a more mainstream technology, with high-end technology of the database will also have better processing ability, that is, the configuration requirements on hardware doesn't have to in the use of some of the high cost of the technology and the equipment. Plays an important role in improve the comprehensive benefit, has become one of the competition focus of the construction, so we are very necessary to grasp the development direction of computer management information system, in order to lay the foundation for the subsequent development and applications. In modern management information system of the design process to fully consider the economic feasibility of the appropriate cost to obtain the interests of the later development of rich returns and the process of economic budget to the initial cost and the subsequent maintenance cost into account. In the figure 5, we show the sample management information system.
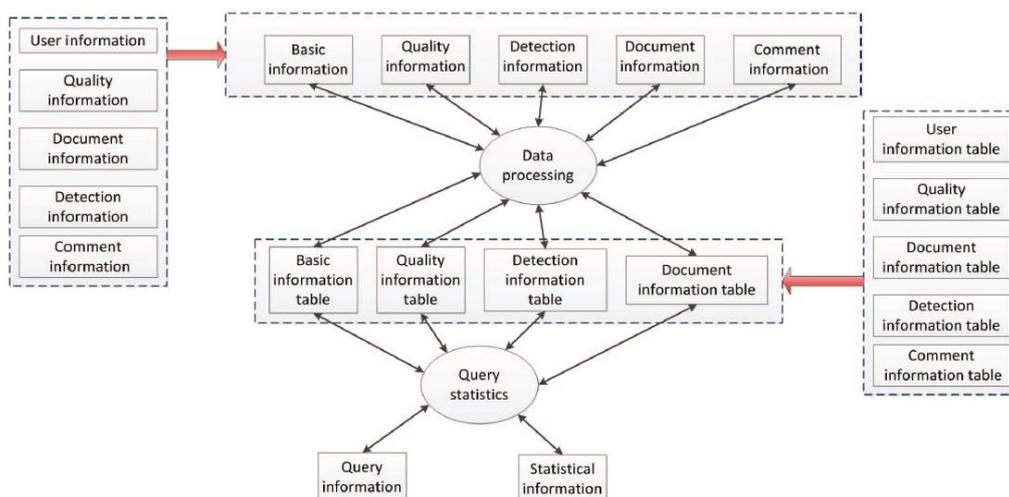

Figure 5. The Sample Traditional Management Information System

4.2. The Modified Novel Management Information System

In modern management, management information is an indispensable intangible resource but also links inside and outside the organization, management and control of important basis and judge decisions necessary foundation, thus the computer management information system development and application of rapidly. We modify the traditional system as the follows. In the figure 6, we show the modified MIS system's flowchart and organization.

(1) Along with the development of communication technology and innovation, and presents the development characteristics of network information processing, prompting to build database and decision support system, and indirectly reflect the computer management information system will be in the direction of network development, and reliance on communication network is more and more big, and this is mainly because must use safe, reliable and efficient information management system of the Internet environment on the management information collection, transmission and processing, and implement different types of data exchange and sharing and processing XML data online, even taking into account the influence of wireless communication, to provide users with different information support and services.

(2) Computer management information system integration is a kind of brand-new concept and method, through to the structure of the complex, dynamic systems and diverse technology as the foundation and support of each subsystem is effective cohesion and integration, in order to complete information integration, function integration, process integration, network integration and enterprise integration,

and ultimately achieve the goal of information exchange and data sharing, and the main important support is based on computer network and the large-scale database.

(3) With management activities becoming more complex, need computer management information system to provide more comprehensive and efficient service, so how to update IT architecture and system, the existing management information system is more flexible and efficient, is a problem urgently to be solved in the present enterprise, and virtual technology based on cloud computing can give full play to its integration server, save money, improve the utilization ratio of resources advantage which is of the great significance to improving the efficiency of the system.

(4) It is widely used in computer management information system at the same time, also inevitable to from the trend of development, has aggravated market competition, enterprises must as soon as possible to make the right decisions, also reflected the system presents the development trend of intelligent, specifically is the future of management information system will be two components, function and process and the process design and process execution separated, thus does not interfere with the process execution in optimization of process design, make the system in the independent component of more rapid processing management information and intelligence technology based on the technology of data mining at the same time, can be found from the management of massive amounts of data timely value data and provide the important basis for decision makers and reference.
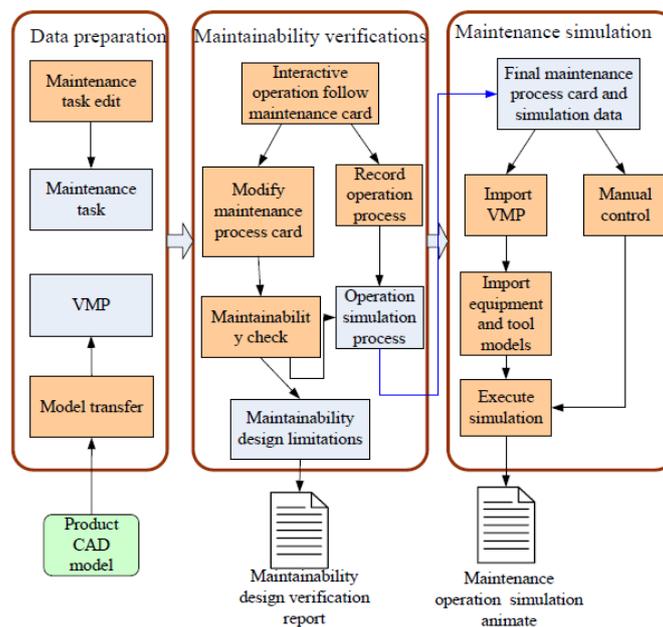

Figure 6. The Modified MIS System for the Service-Oriented Enterprises

5. Conclusion and Summary

In the paper, we analyze service-oriented enterprise management information system based on the large-scale warehouse based optimization and distributed computing methodology. Management information system mainly by the network communication technology, computer hardware and the software technology and with the related support equipment as the foundation, the information collecting, sorting, storage and maintenance, in order to maintain healthy development of the enterprise, benefit and efficiency, to enhance the efficiency of management for the purpose of the development of the management software. Through the use of management information system, the enterprises can solve the allocation of resources and integration problems in operation, and can obtain scientific and objective decision basis, further improve the quality of management of enterprise to enhance the core standardization of enterprise management. Our research obtains advances of data warehouse and the distributed computing to modify the current systems. The numerical performance proves the effectiveness of our design which will help companies manage the resources.

References

Allameh, Sayyed Mohsen, and Sayyed Mohsen Zare. "Examining the impact of KM enablers on knowledge management processes." *Procedia Computer Science* 3 (2011): 1211-1223.

Ball, Marion, Charlotte Weaver, and Joan Kiel, eds. Healthcare information management systems: Cases, strategies, and solutions. Springer Science & Business Media, (2013).

Caniëls, Marjolein CJ, and Ralph JJM Bakens. "The effects of Project Management Information Systems on decision making in a multi project environment." *International Journal of Project Management* 30.2 (2012): 162-175.

Delagenière, Solange, et al. "ISPyB: an information management system for synchrotron macromolecular crystallography." *Bioinformatics* 27.22 (2011): 3186-3192.

Haux, Reinhold, et al. Strategic information management in hospitals: an introduction to hospital information systems. Springer Science & Business Media, (2013).

Holtshouse, D. K. Information technology for knowledge management. Eds. Uwe M. Borghoff, and Remo Pareschi. Springer Science & Business Media, (2013).

Kuperman, Gilad J., Reed M. Gardner, and T. Allan Pryor. HELP: a dynamic hospital information system. Springer Science & Business Media, (2013).

Lee, Seul-Ki, and Jung-Ho Yu. "Success model of project management information system in construction." *Automation in construction* 25 (2012): 82-93.

Majchrzak, Ann, and M. Lynne Markus. "Technology affordances and constraints in management information systems (MIS)." *Encyclopedia of Management Theory,(Ed: E. Kessler), Sage Publications, Forthcoming* (2012).

McKibbon, K. Ann, et al. "The effectiveness of integrated health information technologies across the phases of medication management: a systematic review of randomized controlled trials." *Journal of the American Medical Informatics Association* 19.1 (2012): 22-30.

McPherson, Richard A., and Matthew R. Pincus. Henry's clinical diagnosis and management by laboratory methods. Elsevier Health Sciences, (2011).

Schmidt, Joachim W., and Costantino Thanos, eds. Foundations of knowledge base management: contributions from logic, databases, and artificial intelligence applications. Springer Science & Business Media, (2012).

Tseng, Ming-Lang, Kuo-Jui Wu, and Thi Thoa Nguyen. "Information technology in supply chain management: a case study." *Procedia-Social and Behavioral Sciences* 25 (2011): 257-272.

Wang, Haoxiang, and Jingbin Wang. "An effective image representation method using kernel classification." *Tools with Artificial Intelligence (ICTAI)*, 2014 IEEE 26th International Conference on. IEEE, (2014).

Wang, Yumin, Jiangbo Li, and Harry Haoxiang Wang. "Cluster and cloud computing framework for scientific metrology in flow control." *Cluster Computing* (2017): 1-10.

Xu, Jingjun David, Izak Benbasat, and Ronald T. Cenfetelli. "Integrating service quality with system and information quality: an empirical test in the e-service context." *Mis Quarterly* 37.3 (2013): 777-794.

Xu, Li Da. "Information architecture for supply chain quality management." *International Journal of Production Research* 49.1 (2011): 183-198.

Zaied, Abdel Nasser H. "An integrated success model for evaluating information system in public sectors." *Journal of Emerging Trends in Computing and Information Sciences* 3.6 (2012).