

# Research on the Data Service Application Based on Big Data

Yan Hou

*Qilu Normal University, Jinan, Shandong, 250014, China.*

**Abstract.** The gradual progress of human activities and social production and increasing accumulation scale of information data have become a basic characteristic of modern society. The coming of information era has made the management of information data become the key work of guaranteeing social production and human activity. Especially along with the constant development and maturing of information technology, traditional data management model has not been adapted to the development rate and level of current society. In particular, the wide application of big data technology in all the walks has made the data service application based on big data become the key part of development production. Only by fully considering the data application of the industry in designing big data framework can the visible platform serving the solution be established. This paper will discuss the key skills and strategies of data service application construction based on big data through the analysis on application situation of data service application.

**Key words:** big data; data service application; research.

## INTRODUCTION

In the production activity of society, all the departments should acquire data information through crossing the unit to coordinate the work of each other and guarantee the organized launching. But traditional data storage and handling model have not met with the demand on data application of society. Data information is significant to the manager's solution and strategic plan. Lacking of effective and scientific data information application will badly hinder the production activity of society and influence the long steady development of society. Data information application based on big data should be steady, parallel and effective. It is built on the big data inquiring and analyzing technology, big data storage and handling technology and big data visible technology. The method of data statistics and mining are the common methods to fully ensure the effectiveness of information data acquiring, storage, handling and analysis.

## CURRENT STATUS OF EXISTING DATA SERVICE APPLICATION

### Whole Application Information is Difficult to Obtain

The data expression methods lack the abundance in most current data service application, and characters and graphs are the main means to show the data, which leads to the lacking of effective methods in acquiring the whole application information. As a result, staffs will have the difficulty to information acquiring, analyzing and handling, bad for mastering the whole condition.

### Adaption of Application Demand is Low

Command post is the main direction used in the current data service application, and lacking of various means and flexible models in service guaranteeing and data group leads that it is hard to satisfy the application demand of staffs.

## **Application Range of Data Information is Limited**

Inquiring and statistics of data are the basic ranges of current data application, but the development demand of society must make effective arrangement and analysis with data so as to find the effective data information meeting with the market characteristics and its own development characteristics and provide scientific and reliable data support for decision makers and help them finish the formulation of strategic plan. However, it is very difficult to meet the requirements of decision-makers simply by applying the inquiring and statistical functions of data.

## **SERVICE REQUIREMENTS FOR APPLICATION OF DATA CONSTRUCTION**

The service application and the guarantee command are the core contents of the construction application data, which could effectively integrate and manage the data information obtained from various channels. And data is rich and complex in the information age, presenting a disordered state, which requires the ability to find useful information in accordance with the development and production in a large number of data information in order to facilitate decision-makers to develop targeted and scientific decision-making plans. At the same time, the application personnel should be able to obtain and analyze data at any time according to their own needs, so as to give full play to the maximum value of data application.

1. The application process of operation and maintenance management includes data acquisition and storage, analysis and service. Data security is a key problem in the process of application data construction, so as to ensure that the application data can be more efficient and reasonable in control. In order to meet the need of staff in different levels, a feedback mechanism should be established to scientifically assess and feedback the status of data information application.

2. In order to reasonably divide the resource catalogues of the applied data, the interdepartmental application of data integration and utilization is a problem that should be considered in the process of applying the data construction. The data collection channel should be reasonably standardized to ensure the effective integration of the information data in various departments, which is an important method to ensure the flexible application of the applied data.

3. The application of data service needs to be provided according to the application requirements, which is a key point of the multi-layer decision support in information demand and can ensure no limitation in time and space when we enjoy the application information service support.

4. According to the demand for service guarantee, the data is selected in the complicated and disordered information data, so as to make it convenient for the decision-makers to make the key point of decision-making according to their own needs. It can greatly improve the efficiency and science of data analysis.

## **BASIC CONTENTS OF DATA SERVICE APPLICATION BASED ON BIG DATA**

### **Requirements and Architecture**

Data center should be constructed based on the establishment demand of data, which can play the role of total mastering in integrating data resource and provide data information aiming at the demands of various levels.

### **Support Environment of Big Data**

Whole life period of big data business. It creates the conditions for the establishment of support environment of big data application. It is of great significance to the collection, storage, management and analysis and application of big data. Support environment of big data application can make effective integration with data resource and the scale is rather big so as to satisfy the demand of staffs in different levels.

## **TECHNICAL ARCHITECTURE**

The basic characteristics of the application data are diversity, large capacity and unstructured. In real life, there are different requirements towards the processing and analysis of large data. In order to solve the problem of large data integration, such as sharing, analysis, service and so on, we need to establish a platform architecture to serve and support the life cycle of big data with parallel and efficient features. Data processing and network architecture are the

basic guarantee for the extraction and rational utility of useful information, and the technical framework should work in three aspects: business application, cloud computing platform and large data application support.

1. Establishment of information portal can play a key role in improving the efficiency of finishing business application. “Cloud plus end mode” and “Cloud mode” are the using methods providing for the users so as to make it more effective and convenient in the process of visiting the system.

2. In the practical work demand, the data should be rationally filtered to extract effective data information to start business. Big data service application architecture can be very helpful on integrating and analyzing information and very key in analyzing the development direction of purposes.

3. Technology support of big data is cloud computing platform that can express strong ability in large-scale data computing and storage and exchange, also being the significant technology of guaranteeing data society. On the aspects of change resource, cloud computing platform can also make virtual management with resources.

4. On the basis of cloud computing platform, service support of big data application service provide the collection, storage and analysis. Make the classification and catalogue and input handling of storage., make the direct handling with the Metadata and underlying databases; the searching, extraction and visit can be done with the solution establishment and demand of application demand; multiple expression, whole finding can provide for users with different demands.

## DESIGN PRINCIPLES OF BIG DATA APPLICATION DEVELOPMENT FRAMEWORK

### Business Requirements

Big data processing has the characteristics of wide patterns, fast circulation and high capacity. These characteristics should be fully considered in the process of making application architecture of big data to support the collection, processing and analyzing of big data.

In the application process, big data is required to be expandable, safe and reliable. And such demand characteristics should be taken into consideration in the process of business demand of big data application architecture.

Original data should be carefully integrated and analyzed in the process of application and business of big data application architecture should satisfy its basic requirements.

### Development Framework of Big Data

1. Security, operation and maintenance framework. The security and operation framework of big data are important frameworks for ensuring big data security and maintenance management, including Ambari, Chukwa, Ganglia Context and Embedded Hadoop technology, which respectively play their own key role in arranging, monitoring, managing clusters and data collecting.

2. The connector. The analysis and expression functions of big data can be realized by connecting the database and relation database. ETL is its main technology.

3. Analysis and presentation framework. It is the important framework of data analysis and visibility. Its using process must be with intelligent analysis and presentation tools. Mahout technology can provide data mining algorithm base, and Hama technology can provide scientific computing framework and its scale is rather big.

4. Scheduling framework. Its function is to adjust the big data and launch the analysis work of big data. HBase technology can help to randomly saving and choosing and timely reading in the typing of data; Oozie technology can coordinate with Hadoop work to complete the operation and scheduling of work flow; it also includes Avro and Zookeeper. The integrated development environment provides compilers, automated system builds tools, and source code editors; an integrated application can provide a complete application.

5. Access framework. It is mainly used to visit Hadoop and relation database. Hive, Sqoop and Pig are its main technology. Pig technology can make SQL data analysis on the parallel computing basis of Hadoop; Sqoop can finish the function of data delivery, mainly between traditional database and Hadoop.

6. Processing framework. MapReduce technology can run the application procedure on large-scale cluster system. Owing to its high tolerance for mistakes, it can be perfectly operated when processing the data sets above TB.

7. Storage framework. HDFS technology is usually adopted in the process of storing non-structural data to distribute storage data cluster and combine with the centered management data. It also has the high tolerance for mistakes.

## KEY TECHNOLOGY OF DATA SERVICE APPLICATION BASED ON BIG DATA

### Big Data Storage and Processing Technology

Hadoop technology is widely used in the storage and processing of big data, which can realize its distributed computing and storage. It can help users to improve their work efficiency and work quality when processing large-scale data in the context of the information age. Hadoop technology can easily implement the storage and management of petabyte data. In addition, its extensibility and fault tolerance are obviously superior to other technologies. Hadoop technology mainly includes NoSQL (distributed database), HDFS (distributed file system) and MapReduce (distributed computing framework).

1. NoSQL, namely distributed database. It has a large advantage in high speed reading and writing operation, and it is done in big data volume structured and unstructured data. NoSQL play a key role in high-speed online work.

2. HDFS, namely distributed file system. It can hide the redundant replication and lower load balancing. The API interface is provided for the upper level program. HDFS can optimize the working mode, which is finished on the basis of analyzing its data characteristics. The problem of node failure and the imbalanced ratio of read-write operation is optimized.

3. MapReduce, namely distributed computing framework. It is the abstract operation in distributed storage computing. Key/Value is completed under Reduce and then output the Output.

### Big Data Analysis and Visualization Techniques.

There are some problems to be noticed in big data analysis and visualization technology: traditional online analysis should complete the form shift, and its direction is data finding; structured and unstructured mixed data types replace traditional structured data; it can analyze PB level data; support iterative analysis and figure analysis.

#### 3. Big data query and analysis technology

It is unable to inquire and use Hadoop to store data through SQL. The analysis and operation of big data are required when the staff uses SQL. In order to improve its work efficiency and quality, SQL on Hadoop technology came into being, including Hive, DEAWN Scale, PostgreSQL and so on. Data can be stored in a distributed file system through Hive, and the SQL language can play a key role in the operation of data statistics and analysis.

## CONCLUSION

In the information age, the number and type of information are also constantly expanding, so the construction of data service application is very significant. It is necessary to follow the principle of business requirements and big data development framework, reasonably use big data storage and processing technology, big data analysis and visualization technology, big data query and analysis techniques under the background of big data era to construct the application data.

## REFERENCES

1. Guang Chen. Data Service Application Research Based on Big Data[J/OL]. Computer Technology and Development, 2018(07):1-6
2. Li Hua, Cheng Chen. Meteorological data service application under the cloud computing environment[J]. Agriculture and technology, 2017, 37(20): 231+234.
3. Wenmin Lin. Research on Big Data Service and Key Technologies Under the Cloud Environment[D]. Nanjing university, 2015.
4. Zhikun Chen. Research on the Key Technology of Big Data Organization and Management under the Distributed Environment[D]. University of National Defense Science and Technology, 2014.
5. Yunfeng Liu, Li Li, Sumei Wang, Qianyi Wang, Xu Ynag, Rongbin Ouyang and Xinzheng Long. Design and Application Research of Data Service Center for Big Data[J]. Journal of Communication, 2013, 34(S2): 170-174.
6. Jing Han. Research on the Key Technology of Big Data Service[D]. Beijing University of Posts and Telecommunications, 2013.