

Technology and Application of Large Data Analysis Based on Cloud Computing

Liu Zhongyan, Zhang Lizhe

Nanyang Institute of Technology, Nanyang, Henan, 473000

Keywords: Cloud Computing, Large Data, Analytical Techniques, Applications

Abstract: With the rapid development of modern technology, massive data have been generated, such as micro-blog's message, video data, cell phone terminal, commodity label and so on, which generate a lot of data. Human beings enter the era of big data explosion, that is, the era of big data. The big data age is closely related to cloud computing, Internet of things, and the Internet. In the era of big data, how to complete massive data processing and analyze data in a short time, accurately grasp the potential needs of the target user group, analyze and predict it efficiently, so as to enhance the information technology level of the whole society is very necessary. Cloud computing is the most efficient method for data processing, processing data, improving processing efficiency, and sharing resources.

1. Introduction

At present, the two major development directions of the computer field are big data and cloud computing. Today, in the information age, all kinds of production activities generate a lot of information, which provide the corresponding data basis for big data. The essence of cloud computing is data processing, cloud computing can understand and apply large data, activate data assets, and serve the relevant departments and enterprises. Therefore, cloud computing is the basic technology support of large data, and large data is the development direction of cloud computing. In particular, the value mining of the data can meet the analysis of the report application, and can also benefit directly from the results [1-3]. For example, data analysis can be divided according to the click effect of the system, and transaction data can help them analyze the market and facilitate personalized advertising recommendation.

2. Major Data and an Overview of Cloud Computing

2.1 Definition of data and definition of cloud computing

Cloud computing, popularly speaking, is a technology with super computing power. It can accomplish massive operations that cannot be completed under traditional mode under traditional mode, and achieve prediction of changing trend of things. For example, the forecast of meteorological changes, the trend of market development, the effect of nuclear explosion and so on. A professional term for big data used in data sets is a set of data that laugh beyond the capability of data collection, management and processing data that can be accepted by the commonly used software tools during runtime. The difference between large data and traditional data lies in the ability to store and process massive data. Large data analysis is the ability to optimize the processing of massive data, and has the ability to analyze the problem, and provides appropriate decisions for enterprises or related organizations. There is an essential gap between large data analysis and traditional data analysis, which is more accurate and faster than traditional data processing [4]. Therefore, the main features of the large data are the large amount of data, the complex data sources, and the various types.

2.2 The basic features of large data

The basic features of large data can be summarized as four characteristics: mass, diversity, speed and value.

Today, in the era of massive information, wireless devices, wireless sensors and other devices are

generating data all the time, and hundreds of millions of users are interacting with each other all the time. Scientific research, video surveillance and so on are all important sources of large data. According to the Digital Universe Study report released by IDC in 2011, the global information volume has increased exponentially, doubling almost every two years. Diversity refers to the diversity of data types, which are generally divided into three categories, namely, unstructured, semi-structured and structured data. With the rapid development of science and technology, data types are more complex than traditional data types, such as text, pictures, reports, audio, video and other information are unstructured, and these data account for 80% of the total data [5-6].

Speed refers to the requirements of real-time data processing. In today's big data era, the value of data decreases over time, so real-time, real-time data processing and data analysis, strong interactivity is necessary. Most of the data in the enterprise are structured and managed by relational databases. These data are of high value, but these data are static. However, data from the Internet are mostly unstructured, dynamic and low in value density. How to integrate these two kinds of data is also one of the important features of large data.

3. The Relationship between Large Data and Cloud Computing

Cloud computing provides a corresponding data processing platform for large data analysis, especially large data analysis technology. As we all know, the core of big data analysis technology is data statistics. Data statistics and analysis depend on cloud computing platform. Therefore, cloud computing is the main advantage of big data analysis and processing. Cloud computing provides massive data storage and analysis capabilities, making processing methods more convenient and diversified, and the speed of information processing has corresponding timeliness. The essence of cloud computing is computation [7]. Big data is the object of computation. Big data emphasize more about data storage capacity, while cloud computing focuses on data computing power. The two parties cooperate and cooperate with each other. Single large data statistics can only be simple data acquisition, but can not be applied; single cloud computing, its ability can not be reflected. Therefore, only when the two cooperate with each other, can we improve the value of mutual utilization and promote each other, which has incomparable functions and advantages compared with the traditional information processing and analysis technology, and will become the main direction of future application area.

4. Large Data Technology System

The extraction of data value has been enhanced with the development of cloud computing technology. To dig out more information from the current data, we must face the four basic characteristics of big data. Cloud computing is the foundation of data technology system, and is closely related to big data technology. Big data technology covers massive data storage, processing, English and other aspects of technology, including heterogeneous data sources, massive distributed file systems, parallel computing framework, data mining and data visualization. Therefore, a typical large data processing technology includes data source, data acquisition, data processing, analysis and application.

In the era of big data, Internet and mobile devices have provided a large number of data sources for social enterprises, which is different from the previous internal data, and the difficulty of data collection is also improving. Therefore, in order to collect these data, it needs to be preprocessed, but this is not a unique technology of large data. Secondly, data processing in the era of big data needs to meet six basic characteristics: high expansion, high performance, Gao Rongcuo nature, low cost, easy to use open interface and backward compatibility [8-10]. Data mining mainly includes parallel data mining, search engine / recommendation engine, social network analysis, and so on. Data visualization refers to the process of analyzing the relationship between patterns and data hidden behind data by graphical means, showing the value of data intuitively and effectively.

5. Analysis of Large Data Analysis Technology based on Cloud Computing

5.1 Data mining ability

With the increase of data volume and the diversification of data structure, the development of data mining software has gone through many stages. Now it has entered the era of cloud computing, improving the accuracy and timeliness of data mining. Cloud computing has built a virtual storage system and migration strategy and load balancing system, ensure the effectiveness of data transmission, and realize the unified management for data resource management, parallel tasks in cloud computing, mining guarantee to provide effective and efficient for parallel data. The ultimate goal of data mining is to provide data information for computing processing system. Cloud computing provides segmentation, clustering, outliers and other computing methods can dig deeper into the data, mining the value of the data.

5.2 Semantic association of search engines

The search engine is the process of collecting data from the application system and processing it, and finally providing the user with a process to display the results of the query. A generalized search engine is information retrieval, which organizes information according to certain rules and queries technology according to user needs. A narrow search engine is a process of information retrieval from web, such as Baidu, Google, etc. Big data analysis is to display various kinds of information in various forms, but in practical applications, we need to make comprehensive invocation and further analysis, so that we can fully reflect the value of data [11]. Therefore, we need to carry out semantic standard to facilitate users to extract and process their desired information processing results according to their needs.

5.3 Data visualization effect

Data visualization is to extract massive data by using the processing power of computer image, iconology and technology calculation, and to represent abstract data into visible images and display data association relations, so as to provide corresponding help for relevant institutions and enterprises. Data visualization is applied in the field of non-spatial data. People are no longer limited to relational data, and can view data and structural relationships in an intuitive way. There are a variety of common data visualization models, such as network diagram, tree graph, bubble graph, scatter plot, matrix diagram, bar chart, pie chart and so on.

5.4 Predictive analytical ability

If data quality and data management security are not affirmed and information value can not be displayed, big data analysis technology is applied in management practice through data management and quality realization management. Cloud computing provides corresponding technical support for data prediction and analysis. Data analysts based on their final results, can be rationally analyzed under the support of truth and data mining technology, so as to provide objective judgement for future decisions of related institutions and enterprises [12-13].

6. Big Data Analysis Technology Based on Cloud Computing

6.1 The analysis of large data online

The online analysis combined with a typical case analysis of the application of cloud computing and big data is the data warehouse system. It is with the help of analysis of complex control technology, in the actual decision making full use of the results of the analysis, a process and storage of the resulting visual effect, is convenient for users to query and access. Cloud computing in online can be carried out after the comprehensive analysis of the data, so you can make the decision information to fully feel the data of practical value, so the introduction of OLAP Technology in the data warehouse system, fast and efficient processing of massive data and decision support system for providing a theoretical basis to judge. The main features of the big data is

multidimensional analysis, such as Baidu, Google, Yahoo etc., are stored in the database for statistical information, and data warehouse system integration of commercial terms, by on-line analysis, realizing some search and retrieval with the fastest speed.

6.2 Data mining by big data

Big data analysis technology to the surface in information extraction, can also analyze data, through the tacit knowledge integration to realize information mining association, will be hidden in massive data information in the data mining, the full realization of the practical value of information. The main means of big data analysis of cloud computing technology based on distributed parallel technology, parallel task of cluster splitting in a short period of time, to the computer network in idle equipment are calculated and analyzed with the delivery and for the first time, to a certain extent, improves the data processing speed and analysis ability [14]. Cloud computing provides a data mining model, technical staff can map the Map function to complete the process, improve the efficiency of data mining, and cleaning and processing of independent computing clusters, so as to ensure the final results of data analysis.

6.3 The effect of the application of big data visualization

The visualization application directly reflects the big data analysis ability is big data, it provides data for users of the internal links, can provide the prediction for relevant institutions and enterprises, to provide corresponding theoretical support for the final decision. Big data analysis of cloud computing technology in practical applications based on graphics and image way to provide users with effective visual effect analysis. As everyone knows, the semantic extension of big data, can help users to fully understand the data of chill, improve retrieval speed of information, the breakthrough observed limits provide high reference value to the data analysis results for users.

7. Conclusion

Cloud computing is a hot field of new technology in recent years, permeating all fields of society, guiding human life, work and the change of business model. The development of cloud computing provides corresponding technical support for big data, and can fully tap the value of data. Therefore, cloud computing and big data analysis technology lead to the future development direction of science and technology. Technicians should further study and understand the meaning and technology of the two, combine them organically and apply them fully in practice, conform to the requirements of information age in the era of big data, and improve the actual value of information.

References

- [1] Li J. Analysis of application based on large data and cloud computing platform[J]. Wireless Internet Technology, 2017.
- [2] Yetis Y, Sara R G, Erol B A, et al. Application of Big Data Analytics via Cloud Computing[C]// World Automation Congress. IEEE, 2016:1-5.
- [3] Li Q, Jiang W, Lin S, et al. Examination Data Analysis and Evaluation Platform Based on Cloud Computing[M]// Human Centered Computing. Springer International Publishing, 2016.
- [4] Zhao Y, Calheiros R N, Bailey J, et al. SLA-based profit optimization for resource management of big data analytics-as-a-service platforms in cloud computing environments[C]// IEEE International Conference on Big Data. IEEE, 2017:432-441.
- [5] Liu T, Wu X, Lian Y. Design and construction of sunshine kitchen system based on cloud computing and video data analysis[C]// International Conference on Computer Science and Network Technology. IEEE, 2016:211-215.
- [6] Wang Y J, Yang B, Dai J, et al. Risk Analysis Algorithm and Its Application of Poisonous and Harmful Substance in Food Testing Big Data Based on Cloud Computing[J]. Journal of Anhui

Agricultural Sciences, 2017.

- [7] Chen C. Construction of Big Data Analysis and Decision Support Platform for Library Based on Cloud Computing [J]. Library Theory & Practice, 2016.
- [8] Guo S, Song T, Diao Y. Discussion on the Related Problems of Large Data Analysis Based on Cloud Computing [J]. Journal of Qilu Normal University, 2015.
- [9] Lin H, Zhao J, Jiao Y, et al. Research on building an innovative electric power marketing business application system based on cloud computing and microservices architecture technologies[C]// IEEE, International Conference on Cloud Computing and Big Data Analysis. IEEE, 2017:246-253.
- [10] Wang X. Analysis on data security technology based on cloud computing [J]. Wireless Internet Technology, 2017.
- [11] Yuan Q, Ma C, Lin J. Fine-Grained Access Control for Big Data Based on CP-ABE in Cloud Computing [M]// Intelligent Computation in Big Data Era. Springer Berlin Heidelberg, 2015:344-352.
- [12] Kchaou H, Kechaou Z, Alimi A M. Towards an Offloading Framework based on Big Data Analytics in Mobile Cloud Computing Environments ☆[J]. Procedia Computer Science, 2015, 53(1):292-297.
- [13] Zhang G, Li J, Hao L. Cloud Computing and Its Application in Big Data Processing of Distance Higher Education [J]. International Journal of Emerging Technologies in Learning, 2015, 10(8):55.
- [14] Lin Y. The research of large data processing technology based on Cloud Computing [J]. Electronic Test, 2017.