Research on Sharing System of Big Data Teaching Resources

Zhiqiang Dai [1]

[1] Tourism and Management Project, Jishou University, Zhangjiajie, Hunan, 427000

hunter2011@foxmail.com

Keywords: Big Data; Teaching Resources; Sharing System; Cloud Computing; the Internet of Things

Abstract. Big data and cloud computing and internet of things is the emerging of computer technology, the growth of the data bring new challenges to management. To solve these problems, first described the concept of big data, Secondly analyze the current situation of the development of big data, large data is given in the application of teaching management, Finally proposed the construction of big data of teaching resources sharing system key technologies.

Introduction

Big Data is one of the hot topics in the next decade will be the era of big data leading advanced technology, regardless of whether people are aware of its existence, big data will transcend and change people's way of life. But the exact definition of big data, it has not been given in the past with respect to local, small data, big data scale is massive. • Victor Meyer - one of the data scientists insight into trends Schonberg was the first big data era, he used to describe the three characteristics of large data [1]:

1, In the era of large data, one can analyze more data, and sometimes can even handle all the data associated with a particular phenomenon, rather than rely on random sampling. For example, a study of sumo competitions rigged results of the competition on 64 000 games were analyzed, this is not really a big figure, but because this is the last decade all the games, so it is big data.

2, Thanks to more data, people can accept more mixed, more imprecise data on. For one thing, if people only 50 data points, then each data point must be very precise, because each data point is useful; but if there are 50 million people, get rid of 10, or even get rid of 1000 are not too big problem.

3, No longer elusive search for a causal relationship, and focus instead on the correlation between things. Analysis of large data primarily in order to predict the future "is what" instead of "Why." Because a lot of times people believe they have found the reason behind things, in fact, did not find. For example, know where the flu will spread to enough, people do not need to know why; know when to purchase tickets online can get the best price on enough people I do not know why this time the lowest price.

Large data volume of data is huge, rapid growth in the size of the data at the same time, the value of the data generated will also be simultaneously increased, some people think that the value of data will increase as data is reduced, this happens mainly because the existing technology can not meet the needs of data growth, the lack of hidden information from large data mining in technology and methods. When you earn nine dollars from 1GB of data, and give you 1PB of data, you can only earn 15 dollars. And if it is proportional to the value and size of the data, then the 1PB of data can give you $ 9 million. For really large data, increase the value of their data should be proportional to the growth, even faster.

Big Data Development Status

Businesses government agencies, academia, industry and so on have already begun to pay close attention to the big data problem, and its produce deep interest, "Nature" in 2008 on the launch of Big Data magazine, made mankind has entered the era of petabyte [2], and predict the next IT giant's main business will be the big data management. "The fourth paradigm" Microsoft Research

published a book explains how to do data-intensive scientific research in the e-Science era. "Science" in February 2011 launched the magazine "Dealing with Data" [3], discussed in detail the development of big data problem. US well-known experts and scholars in the field of data management issued a White Paper on "Challenges and Opportunities with Big Data", the white paper [4] from the introduction of a large data analysis of large data processing, and gives big data challenge. World-renowned consulting firm McKinsey in June 2011 issued a report on Big Data "Big data: The next frontier for innovation, competition, and productivity", the key technology and application of environmental and other big data for a detailed analysis and described [5]. In March 2012 the Obama administration issued a "Big Data Research and Development Initiative," the official launch of "big data development plan." This program is regarded as the Obama administration after the US government information to adjust the road (Information Highway) program is another major initiative in the field of information science.

Chinese scholars are also closely related to big data technology, in 2012 Hadoop and Big Data Technology Conference to "big data sharing and open technology" as the theme, set the "Hadoop", "big data industry application", "big data sharing platforms and applications, "the five sub-forums outlook bright future development of large data. On the 16th National Symposium on regional tourism development, the paper "smart tourism and its key technologies and framework research" describes big data, networking, cloud computing and other new generation of information technology in the wisdom of Tourism and Construction wisdom framework of tourism.

Big Data Applications in Teaching

Our hospital invested several hundred million yuan to complete the design of cloud computing platform, set up a joint enterprise software engineering training center, including software development sub-chamber, cloud computing and networking sub-compartment room. The future will be big data research and scientific innovation platform, to achieve full liberalization of the inside and outside of teaching resources, open content including hardware and software facilities, computing resources, educational resources, content and other resources, experiments and experimental teaching system and teaching model, strengthen business school or research institutes to cooperate.

Software Engineering Training Center, College of Computer software and hardware configuration is complete, the system safe and C/C ++, FORTRAN, CUDA C compiler system, math library and MPI libraries. System installed Hadoop, HBase, ABINIT, CPMD, NAMD, GROMACS, LAMMPS, BLAST, mpiBLAST, WRF, MM5, Graphs and other application software, calculations involving physics, computational chemistry, molecular dynamics, sequence alignment, molecular docking, structural mechanics , structural dynamics, fluid mechanics, mesoscale simulation. The center realized by sharing teaching resources, teaching first to collect data, the data into data assets teaching, teaching collect data to be fine, to be exact, in fact, is ordered, the data needs to be stored, valuable data changes , flowing, and other data interactions. For teaching shared, the best system is to establish a basis for resource sharing large data platform, to achieve full liberalization and sharing of teaching resources.

In teaching resource sharing platform is the most important part is the data processing layer, in the data processing layer, parallel load data storage module to become the core of the platform to achieve, Hadoop distributed technology provides a platform for data storage and data processing models and methods.

Using the Hadoop Distributed File System source data store massive amounts of teaching.

Through the MapReduce distributed computing model to deal with these massive source data.

Using HBase distributed database processing mass data storage, storage management in order to achieve the development of mass education and sharing of data.

Conclusion

With the rapid development of computer technology and the dramatic growth of data, people are surrounded by huge amounts of data, the era of big data has arrived. Rational use of technology to manage large data teaching resources to achieve full sharing of teaching resources is bound to bring great convenience to the management of unversities, but also brought challenges. In this paper, the concept, development status and other big data are described in detail. It describes the characteristics of big data, analyzes the key technology utilizes to big data sharing teaching resources, but also encounters many problems in the design process, such as data integration, management and other issues. Solve big data problems still need to go a long way and this is the future direction of our continuing research.

References

[1] Victor, Meyer, Schonberg. Big Data Era. Zhejiang People's Publishing House, 2013-01-01

[2] Nature. Big Data[EB/OL]. Http:// www.Nature.Com/News/Specials/Bigdata/Index.Html.

[3] Science. Special Online Collection: Dealing with Data [EB/OL]. Http://Www. Science Mag. Org/ Site/ Special/ Data/.

[4] Agrawal D, Bernstein P, Bertino E, Et Al. Challenges And Opportunities With Big Data-A Community White Paper Developed by Leading Researchers Across the United States [R/OL]. Http:// Cra.Org/Ccc/Docs/Init/Bigdata Whitepater. Pdf.

[5] Manyika J, Chui M, Brown B, Et Al. Big Data: The Next Frontier for Innovation, Competition, and Productivity [R/OL]. Http://www. Mckinsey. Com/ Insights/ MGI/ Research/ Technology and Innovation/ Big Data the Next Frontier for Innovation.