

Research on The Network Cloud Teaching Mode of PBL Project Teaching Method

Lei Zhang¹, Zheng Ren², Xianming Shi¹, Qingqi Zeng^{1,3} and Chao Nie¹

¹Jiangsu Jiankang Vocational College, Nanjing, 210029 China.

²The Second Middle School of Zhenjiang, Zhenjiang, 212004, China

³Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, 210029, China

jsjkc@hotmail.com

Abstract

This paper designed the learning resource storage model mainly based on Hadoop platform, and the system structure was optimized on the basis of the original system against the problems of inefficient storage of massive small files, high memory overhead on the Hadoop platform. The PBL teaching resources are classified to improve the utilization ratio of resources and access efficiency of resources according to the file size and file correlation. When processing the small files, small files are merged according to the knowledge point, and then storage processing is conducted with the method of handling large files, giving full play to the advantages of processing files in the Hadoop. Simulation experiment is conducted in the laboratory environment, to compare the integrated storage space with the time required for resources storage. The result is that the model can reduce the storage space of the file, and improve the integrated storage efficiency on the basis of not affecting file storage.

Keywords: PBL teaching; Massive resources; Learning resource; Cloud storage

1. Introduction

As information technology, multimedia technology unceasing infiltrate into education field, the forms of digital learning resources become increasingly complex and contents get constantly enriched. Combined with the application of network in various fields in society, and the dramatic increase in Internet users, online learning has gradually become the mainstream learning form, and the network is the major means and the way of daily access to learning for learners. Higher requirements of resources storage and service are put forward in this situation. Traditional learning repository is used to solve such problems, and the method of upgrading the hardware is usually adopted, causing very high update frequency of hardware equipment, generating a large number of idle equipment, and waste of resources. At the same time, with the constant growth of learning resources, there will be hardware bottleneck. Thus, PBL teaching method integrates the traditional LBL teaching method, which can effectively improve the teaching quality, and improve the comprehensive quality of students. However, there still exist some problems in the process of implementation. The biggest problem is insufficient teaching hours for PBL, and PBL teaching method is only used for a few of selected key and difficult problems in order to save class hours. In addition, the hybrid teaching method needs teachers to have strong comprehensive quality and classroom control ability; Part of students with poor autonomous learning ability are not easy to adapt to the PBL teaching method, they need to the detailed guidance of teachers, so the higher requirements for teachers are also put forward, and the teaching workload is increased. Therefore, we will further refine the implementation plan of hybrid teaching method, achieve the most

effective fusion of two teaching modes and expand teacher team, to let more teachers and students participate in it, and benefit from it.

2. Learning Resource Storage Model Structure and Function Design

2.1. Learning Resources Analysis

Because network technology, computer technology and multimedia technology are applied to the education field, the digital form of learning resources is increasingly complex. It is not a simple picture, and a text, but the multiple resources of media forms constitute a complex. A good learning resource takes a knowledge point as the link, the text, images, audio, and video and other multimedia resources are put together according to a learning process. These composition elements are not isolated but interrelated. So these files storage is not like simple storage files, but the relationship between them needs to be considered. Only considering this can these learning resources be put together better, to provide good service for learners. In addition, the digital learning resources contain a large number of video, audio, images, text and other forms of files. These resources are accumulated over a long period in the development process of each discipline, and have become massive data resources, which puts forward new requirements for the storage device. Now the form of large capacity disc or disk array is generally adopted, which has higher requirements for hardware.

2.2. Learning Resource Storage Model Structure Design

Hadoop file system uses the principal and subordinate structure of the single node, this structure can guarantee the consistency of resources well, prone to the effective management of resources, but this structure puts great pressure on the center node and easily creates bottlenecks, therefore, this study adopts the principal and subordinate structure of the single node, and optimizes its structure. The concrete structure is shown in Figure 1.

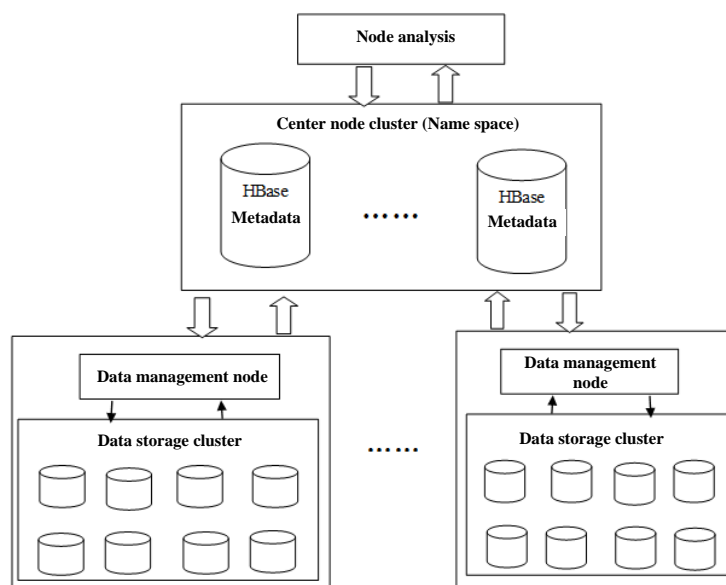


Figure 1. Overall Resources Structure Based on Hadoop

✓ Analysis node

Analysis node is the service node used to control the name node, and the analysis is carried on to determine which name server to provide services according to the information submitted by the users. To provide service for the user, and store the relevant information of name node is an important node in the model, and the storage and access of the resource file are controlled by the lower name node cluster. When analyzing the name node resource allocation information related to the node storage, after the user submits the command, the simple query is carried on, to gain the name node path of the access, and pass the task to the corresponding node name. Based on the analysis of less node calculation, adopting single machine won't produce bottleneck and fault, but considering the stability of the system, it is recommended to adopt dual-machine backup during deployment. When there is any abnormality on the master node affecting the services, satellite computer responds immediately, and replaces the computer to complete the task, ensuring that the task is completed smoothly, so as to further improve the stability of the system.

✓ Name cluster

The name cluster in the architecture is a server cluster used to store the learning resource file metadata. Database uses the own HBase of Hadoop system to complete. Taking the name, type of learning resources and the metadata as the key value <Key,Type,Value> can avoid the occurrence of repeated resources.

✓ Data control node

Data control node is low-end control unit of data node cluster, responsible for heartbeat detection, load balancing and cluster management of the cluster, similar to the center node in the traditional Hadoop, but it does not store the metadata information.

✓ Data cluster

Data cluster is composed of a large number of computers with low performance, used to store file block information of learning resources, communicates with the data control node in real time, and feeds back the heartbeat detection information. When the node encounters problems, the data control node is responsible for allocating and scheduling the new node to provide service. Learning resources are partitioned according to the certain rules, and multi-copy storage is conducted on these file blocks on the cluster by distributed storage strategy.

2.3. PBL Teaching Process

PBL teaching is the process of causing the independent learning of the students according to the problems, its main process includes seven stages of situation problem, team building, technical thought, data collection, theoretical study, scheme realization and examination and evaluation and the specific implementation process is shown in Figure 1.

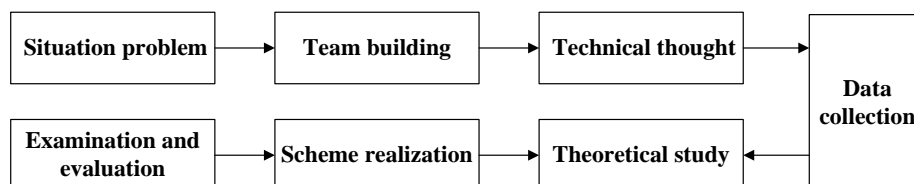


Figure 2. Implementation Process of the PBL Teaching

(1) Situation problem: Situation problem is the center of the PBL teaching, and determines the effect of learning to a large extent. We design the situation problems with

appropriate design difficulty and wider knowledge coverage and combining the actual application for the specific questions of important and difficult points and in-class experiment and large-scale experiment.

(2) Team building: there is one to three learning team personnel, allow independent group, and encourage team members to match with the strong and the weak.

(3) Technical thought: when PBL teaching is used for the learning of key and difficult problems, technical thought will be mainly provided by the teachers, LBL teaching method is adopted to explain theoretical knowledge and key technology involved in detail, and the corresponding resources website is provided. When the PBL teaching method is used for in-class experiment and large-scale experiment, technical thought must be built by the students, and teachers only provide consultations and guidance.

(4) Data collection: students collect the data that theoretical study needs based on the guidance of teachers, with the help of Internet, library and other platforms.

(5) Theoretical study: students carry out the theoretical study and discussion on the collected information under the teacher's explanation and guidance, complete the design of the scheme (including class design, algorithm process, *etc.*,) and form the corresponding technical documentation.

(6) Scheme realization: students realize the scheme, solve the problem and submit the document through programming, debugging and testing. The teachers provide document template in the process, and guide students to write the technical documents according with the engineering specification.

(7) Examination and evaluation: examination and evaluation are conducted by teachers according to the students' achievement and are incorporated into the performance record, and the common problems and errors of students are summarized, and detailed explanation and feedback are conducted with LBL teaching method.

2.4. Distributed Storage Design of Learning Resources

Learning resources generally include video, sound, pictures, literature, and courseware, *etc.*, due to the continuous development of multimedia technology, these resources increase, and some single file size is greater than 1 GB (such as video files), only a few KB (such as text data and image data). These files can't be stored like ordinary file, and there is the internal connection between these files in the massive data storage, so we fully consider the connection between files when designing the distributed storage file to store the file with the same knowledge into a node in order to reduce the document retrieval and read-write access. Because Hadoop platform will cause the memory overhead and the increase of read-write time in processing massive small files, but it has obvious advantages to the large file storage. According to different file sizes, the storage time of Hadoop is as shown in Table 1.

Table 1. Storage Test Results of Various Files

File size/MB	2MB	4MB	8MB	16MB	32MB	64MB
Execution time/MS	1404	1021	850	700	589	303

Through the analysis of the execution time, when the file size is close to 64 MB, running time decreases gradually. According to the characteristics of the Hadoop, all the learning resource files are turned into the large files for storage and processing. The file is divided into large and small files when processing the file, and two types of files adopt different processing methods. When considering both the execution time of files and the generation of disk fragmentation of files, below 64 MB is defined as the small files, and more than 64 MB is defined as the large files.

2.4.1. Large File Processing: The large files are stored in blocks, each block is 64 MB, and all blocks belonging to the same large file are stored as a file, and the distributed storage scheduling policy is used to store all file blocks on the node separately, and the corresponding health degree of BT file seed [4] is set up. In order to facilitate retrieval, data block of each file adopts the form of < ID, Data > for storage, of which ID is the serial number identification of the block, and the Data is the data content, so the file can be quickly retrieved through ID number.

2.4.2. Small File Processing: The small files shall be merged into one file queue in accordance with the internal relations, the file queue is taken as a large file, and the large file storage method still adopt the traditional method in the Hadoop for processing. When creating the file sequence, we consider using the knowledge point for partitioning, and the same or similar knowledge point resource files are merged by the keyword, convenient for the access of users to access. The queue structure is as shown in Figure 2. Size is used to store the file size of current block, easy to continue to add on the original file later, Knowledge_Key is to store the key words of the knowledge points, Key uses the file name as the identification to number the file information, and value is used to store the file content.

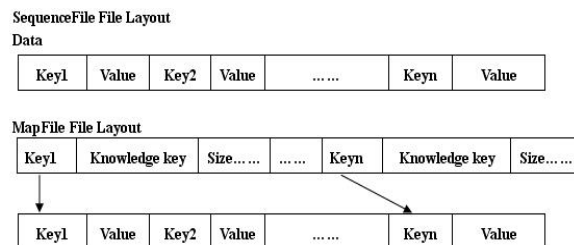


Figure 3. Storage Structure of Small Files

2.4.3. Learning Resource File Metadata Storage: Most of traditional learning resources storage uses relational database for storage. When the relational databases deal with massive data, service efficiency (including query, insert, *etc.*) will decrease, and system bottleneck can easily arise. In order to meet the requirements of massive learning resources storage, HBase in the Hadoop is used to store the metadata of resource. Non-relational database storage form can divide the table, to store different tables into different nodes, and it has good extensibility.

According to the analysis of learning resources, each resource file has its own independent filename, the file name and the storage time are combined as the Key value, and other properties are described as the column family. The files with internal relation are stored nearby, to reduce the access time of reading the file by the storage devices. Metadata is stored with the structure as shown in Table 2. RowKey is the only identification of the metadata file, the combination of time and resource file name is adopted in design, and what is the same batch of the submitted resources file can be determined by time, convenient for the realization of the nearby storage. ColumnFamily is column family, and is divided into many columns again, to describe the file, including the file size, type, block Id, keyword, name, *etc.*, and it can be extended according to specific needs.

Table 2. Learning Resource File Metadata Storage Structure

Row Key	tamp	Column Family			
		Size	Type	Name	Keyword
20C29	T1	16K	Jpg	Jpeg1	Flow chart
	T2	105K	Doc	Do1	Program
20C2f3	T5	96K	Xls	Book1	Football

3. Simulation Experiment Analysis

This research conducts the system simulation test under the laboratory environment. Model simulation is conducted on the seven computers, one computer is set as the control node, two computers simulate name space nodes, four computers simulate data management nodes, and four Linux operating systems are installed on the virtual machine used by the data management node, to simulate data cluster. The file input is implemented on the control node through the Tomcat service program, Hadoop center node program and database program are deployed on the name space nodes, and cluster node program is installed on the data control node, forming a Hadoop platform. Main test system conducts the space cost in storage and the time cost in storage on the learning resource files, and each test uses the average value of 10 tests.

3.1. Storage Space Test

100 large files of more than 64 MB and small files between a few KB to 4 MB are adopted to conduct the storage test. The Hadoop default storage and storage under this model are respectively used to store the files in turn according to the file size, and the occupied space is compared shown in Figure 3. By contrast, we find that when Hadoop processes the file, storage space will increase sharply with the increase of small files, and the storage space in this model is superior to the traditional Hadoop storage mode. Integrated file space = large file space + small file space.

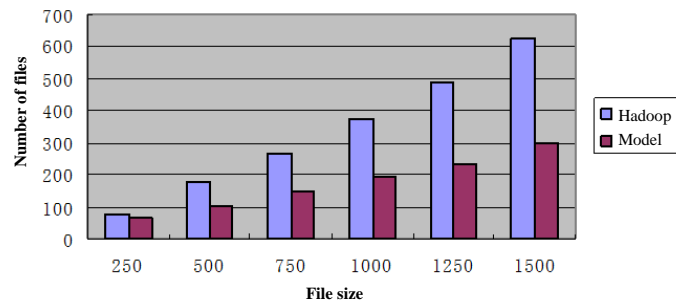


Figure 4. Files Storage Space

3.2. Storage Time Test

According to the number of files, the files are grouped into 10, 100, 200... 1000, 1100 files when testing, and the storage time is as shown in Figure 4. According to the time-consuming test, we find that when Hadoop stores files, time-consuming increase sharply with the increase of small files. The storage time of files in the model is basically constant.

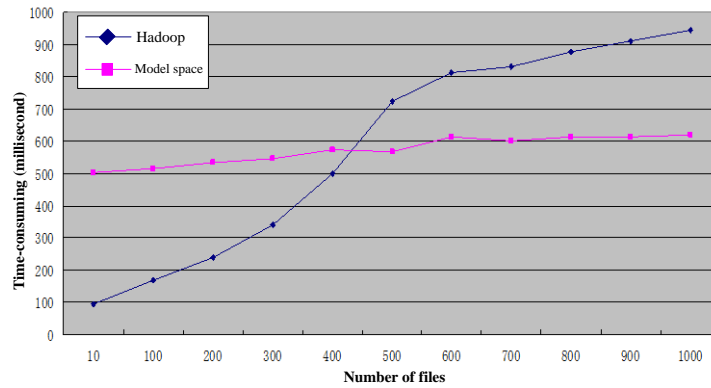


Figure 5. Time-Consuming for Storing Files

4. Conclusion

This paper optimized the storage architecture of file system on the basis of Hadoop, and designed the storage structure for learning resources file distributed storage. The resource files are classified for storage in accordance with the characteristics of the learning resource files, and the files with the internal relations are stored nearby, which solves the hardware bottleneck of resource file storage, improves the utilization rate of hardware resource and raises the access efficiency of files. The application of cloud computing ideal in the education field is worth researching.

Acknowledgment

The research is supported by Key projects of the reform of the Education Department of Jiangsu Province, China under Grant No.2011JSJG071.

References

- [1] J. Yang, B. Chen and J. Zhou, "A Low-Power and Portable Biomedical Device for Respiratory Monitoring with a Stable Power Source", *Sensors*, vol. 15, no. 8, (2015), pp. 19618-19632.
- [2] G. Bao, L. Mi, Y. Geng and K. Pahlavan, "A computer vision based speed estimation technique for localizing the wireless capsule endoscope inside small intestine", 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), (2014) August.
- [3] X. Song and Y. Geng, "Distributed community detection optimization algorithm for complex networks", *Journal of Networks*, vol. 9, no. 10, (2014) January, pp. 2758-2765.
- [4] D. Jiang, X. Ying and Y. Han, "Collaborative multi-hop routing in cognitive wireless networks", *Wireless Personal Communications*, (2015), pp. 1-23.
- [5] J. Hu and Z. Gao, "Modules identification in gene positive networks of hepatocellular carcinoma using Pearson agglomerative method and Pearson cohesion coupling modularity", *Journal of Applied Mathematics*, vol. 2012, (2012).
- [6] D. Jiang, Z. Xu and Z. Chen, "Joint time-frequency sparse estimation of large-scale network traffic", *Computer Networks*, vol. 55, no. 15, (2011), pp. 3533-3547.
- [7] J. Hu, Z. Gao and W. Pan, "Multiangle Social Network Recommendation Algorithms and Similarity Network Evaluation", *Journal of Applied Mathematics*, vol. 2013, (2013).
- [8] M. Zhou, G. Bao, Y. Geng, B. Alkandari and X. Li, "Polyp detection and radius measurement in small intestine using video capsule endoscopy", 2014 7th International Conference on Biomedical Engineering and Informatics (BMEI), (2014) October.
- [9] G. Yan, Y. Lv, Q. Wang and Y. Geng, "Routing algorithm based on delay rate in wireless cognitive radio network", *Journal of Networks*, vol. 9, no. 4, (2014) January, pp. 948-955.
- [10] Y. Lin, J. Yang and Z. Lv, "A Self-Assessment Stereo Capture Model Applicable to the Internet of Things", *Sensors*, vol. 15, no. 8, (2015), pp. 20925-20944.
- [11] K. Wang, X. Zhou and T. Li, "Optimizing load balancing and data-locality with data-aware scheduling", *Big Data (Big Data)*, 2014 IEEE International Conference on. IEEE, (2014), pp. 119-128.
- [12] L. Zhang, B. He and J. Sun, "Double Image Multi-Encryption Algorithm Based on Fractional Chaotic Time Series", *Journal of Computational and Theoretical Nanoscience*, vol. 12, (2015), pp. 1-7.

- [12] T. Su, Z. Lv and S. Gao, "3d seabed: 3d modeling and visualization platform for the seabed", Multimedia and Expo Workshops (ICMEW), 2014 IEEE International Conference on. IEEE, (2014), pp. 1-6.
- [13] Y. Geng, J. Chen, R. Fu, G. Bao and K. Pahlavan, "Enlighten wearable physiological monitoring systems: On-body rf characteristics based human motion classification using a support vector machine", IEEE transactions on mobile computing, vol. 1, no. 1, (2015) April, pp. 1-15.
- [14] Z. Lv, A. Halawani and S. Feng, "Multimodal hand and foot gesture interaction for handheld devices", ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 11, no. 1, (2014), pp. 10.
- [15] G. Liu, Y. Gen and K. Pahlavan, "Effects of calibration RFID tags on performance of inertial navigation in indoor environment", 2015 International Conference on Computing, Networking and Communications (ICNC), (2015) February.
- [16] J. He, Y. Geng, Y. Wan, S. Li and K. Pahlavan, "A cyber physical test-bed for virtualization of RF access environment for body sensor network", IEEE Sensor Journal, vol. 13, no. 10, (2013) October, pp. 3826-3836.

Authors



Zhang Lei, received her M.S. in Nanjing Medical University, China. She is currently an associate professor in Jiangsu Jiankang Vocational College at Nanjing. She is interested in network technology. Also, she has got several related certificates. Her research interest is mainly in the area of education and teaching reform. She has published a lot of research papers in scholarly journals in the above research areas and has participated in a number of teaching materials. E-mail: jsjkzl@outlook.com



Nie Chao, received his M.S. at Nanjing University of Chinese Medicine. Nowadays, he is a senior lecturer in Jiangsu Jiankang Vocational College in Nanjing, China. He focuses on Computer Software, especially the computer network technology, cloud technologies and so on. He has designed and programmed a fraction of routines and has got the certificate of level three in National Computer Rank Examination in 2002. In addition, he has published a lot of papers in scholarly journals in the area of Teaching Reform. E-mail: jsjknc@hotmail.com