Research on the Construction of Convergence Media Based on Big Data Technology

Qiru Zi(✉)

Department of Media and Communication, Mianyang Teachers' College, Mianyang, China
ziqiru113@gmail.com

Abstract. Big data technology, as one of the representative new technologies in the era of science and technology, is widely used in various fields. This paper focuses on the application of big data technology in the construction of convergence media, explains the main advantages of big data technology in detail, and discusses the main problems that need to be faced in the construction of convergence media one by one, summarizes relevant experience and gives targeted development suggestions, hoping to provide reasonable reference for workers in the same field.

Keywords: convergence media · big data technology · construction · data collection · data processing · cloud computing

1 Introduction

Under the background of science and technology era, the volume of user data in the network is increasing, and people have higher requirements for the use of existing radio and television engineering technology and data processing effect. As a result, the convergence media industry has emerged in the public's field of vision, and people's traditional daily life habits have changed greatly. At present, with the continuous development of big data technology, the convergence media radio and television engineering technology has made great progress in data analysis and data processing, and it can effectively screen and integrate massive data content with the help of convergence media and internet technologies, so as to achieve a high-quality combination of interactivity and differences, which provides a new development direction for the radio and television media industry. Based on this, in order to further meet people's actual demand for the media industry, in the subsequent development process, it is necessary to continuously strengthen the application effect of big data technology, and rationally use big data technology in the convergence media radio and television engineering technology, thus laying a better environmental foundation for the future development of convergence media radio and television engineering.

2 Overview of Big Data Technology

Big data technology is a key component in the field of modern intelligence. Under the background of big data, with the changes of people's existing life and work mode, intelligent services have become more common in people's lives and become the mainstream

K. Hemachandran et al. (Eds.): ICAID 2023, AHIS 9, pp. 348–357, 2023.
https://doi.org/10.2991/978-94-6463-222-4_37

development trend of modern service methods, such as smart shopping and intelligent navigation. Electronic information technology can bring better automation and intelligent services to people from these perspectives [1]. In the era of big data and the continuous development of cloud computing, by further mining key data information, we can not only save human resources more effectively, but also closely combine the practical needs of users with information technology, and make full use of data acquisition, data analysis, data mining, cloud computing and other ways to optimize and integrate the information resources sorted out by electronic information technology, so as to accurately find out the data with high value, and the application effect is excellent.

3 Application of Big Data Technology in the Construction of Convergence Media

3.1 Data Acquisition

Big data technology is the fundamental technology in the information technology of convergence media radio and television engineering, and data information belongs to the core content of building a modern convergence media platform. In order to ensure the standardization and standardization of data collected by the system, the best treatment measures are: centralized deployment of a unified management platform. The unified management platform here mainly refers to the full-process business service platform formed by highly integrated technologies. The system can provide the following main service functions: intelligent voice recognition function, business management function and command and coordination function. During this period, the business management function needs to ensure that all business links are clearly presented and recorded, and that all data collected by the system are normative, standard and accurate. After the corresponding instruction information is accurately entered in the Internet system, some Internet of Things systems have not set a unified format standard for trace management data, which leads to the failure to achieve unified coordination among different Internet of Things systems and timely coordination and cooperation among individual Internet of Things systems to a certain extent, but they still do not have a unified public foundation [2].

In the above content, during the completion of the data information collection task, the existing convergence media radio and television engineering technology needs to complete the basic information collection task with the help of network technology, and at the same time, it can also collect and sort out the data in Weibo or other forums in a centralized way, or dig out deeper representative data information with the help of these data resources. After that, with the help of background monitoring, we can control all the data information collected by ourselves in real time and ensure the timeliness of real-time dynamic update. Under such circumstances, we can greatly improve the work efficiency of traditional convergence media radio and television engineering technology in data collection and data collation.

In addition to the above contents, we can also use big data technology to complete the collection of picture information and video dynamic information, and ensure the

authenticity and richness of the data mastered by the convergence media radio and television engineering technology. Therefore, it is necessary to use information technology to upload data content to the cloud and save it, so as to ensure its own data security.

3.2 Data Analysis

Under normal circumstances, data information has the characteristics of diversity, so it is necessary to make a comprehensive analysis of data content with the help of professional network technology in order to make a deeper understanding of users' real needs. On this basis, we can take more targeted measures to provide users with better comprehensive services, and inject new impetus into the future development of the existing radio and television media industry to promote its continuous development in the service direction [3]. During this period, because the informatization construction of data resources needs to be promoted in a top-down order, in the actual operation process, it will be influenced by staff, such as the age factor of personnel, personal knowledge system, personal work nature and many other elements, which leads to great variables in some media informatization construction work, especially in the context of big data era, some personnel do not have enough unified system design in data management. Without the corresponding practical management experience, these staff members themselves lack a complete and correct awareness of big data, and most of them rely on their own intuition and experience in the process of completing data management tasks, which will also affect their own information construction level. Figure 1 shows a schematic diagram of the public cloud technology architecture provided by the Internet.

In addition, due to the lack of reliable security technology for real-time information in the Internet system itself, under special circumstances, the problem of untimely data sharing is more serious, which will also lead to differences in related contents included in the database, and can not guarantee the accuracy of relevant implementation standards and key data. Finally, it affects the normal function of service and can't fully show the

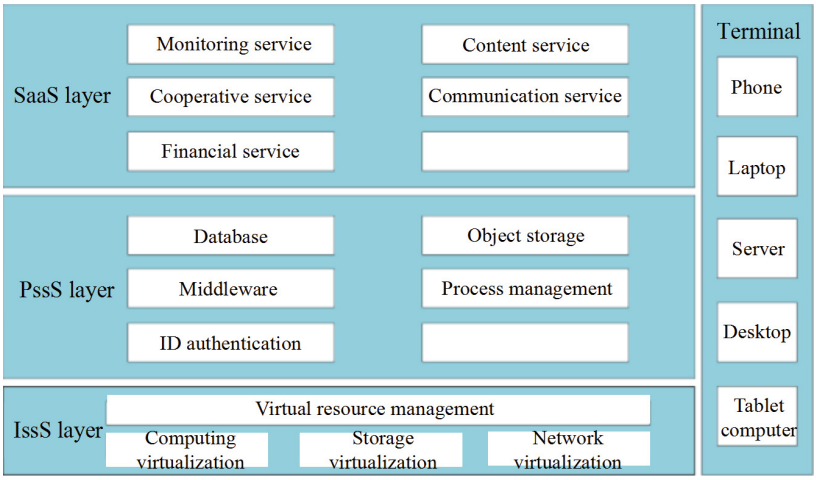

Fig. 1. Schematic diagram of public cloud technology architecture provided by the Internet

actual use effect of nursing network information system. Some Internet information systems have technical problems, and there is no targeted analysis technology for the data stored in the system. Therefore, in the face of the massive content in the database, it is impossible to comprehensively process all the content through big data technology [4]. Based on this, it is necessary to further improve the service ability of existing system decision-making, ensure the final application effect of big data technology, achieve the goal of improving service quality in this way, and fully demonstrate the information advantages of big data technology.

4 Application of Big Data Technology in the Construction of Convergence Media Platform

Based on Hadoop and Spark technology, an information analysis system platform can be established. The main contents are as follows:

(1) Data source layer: The data in this system module can be divided into two forms, the first is the real-time data fed back by network users in daily life, and the second is the data provided by static business system. During this period, there will be various forms of historical data, personal basic information, file information, etc., which are classified as static and unchangeable types of knowledge data. During this period, with the passage of time, the accumulation of various data will lead to a significant decrease in the actual update speed of the system [5].

For the above knowledge data content, it includes not only the traditional relational database, but also a lot of video information, audio information, text information and other forms of information. Most of these unstructured data will be stored in the platform system in the form of structured or unstructured existence. Table 1 shows the Public cloud SaaS layer service information table.

(2) Data acquisition and management layer: This control layer is mainly responsible for extracting the data provided by the data source and transferring it to the platform system. Among them, with the difference of data sources, the final collection methods of data information are quite different. Take the content of traditional relational databases such as MySQL, SQL Server and Oracle as an example, and its data needs to be imported by Sqoop. The overall service architecture design of broadcast private cloud is shown in Fig. 2.

Table 1. Public cloud SaaS layer service information table

	Configuration	High-performance	Extendable
Level 1	N	N	N
Level2	Y	N	N
Level3	Y	Y	N
Level4	Y	Y	Y

(3) The data processing module needs to centrally process the data standardization, and in the process of data processing, it also needs to remove duplicate data, missing data and other types of unavailable data in time. After that, the missing data can be imported again, and the corresponding processing tasks can be completed by manual entry or logical automatic completion. For the processing of some erroneous data, such as abnormal data processing, the approximate value of data can be obtained by removing or based on historical data with the help of interval definition. For some duplicate data contents, such as all duplicates, the duplicate contents can be removed, or they can be selectively removed according to the time sequence or business logic at different stages. For some unusable data, you can also use the rule adaptation processing method, keyword matching processing method and enumeration transformation processing method to complete the necessary repair task or cleaning task.

(4) Data analysis layer: This system module needs to use the Spark computing framework to complete the analysis of key data. Among them, when Spark reads the data content in Hdfs, Hive or HBase, it can create the corresponding RDD. After that,

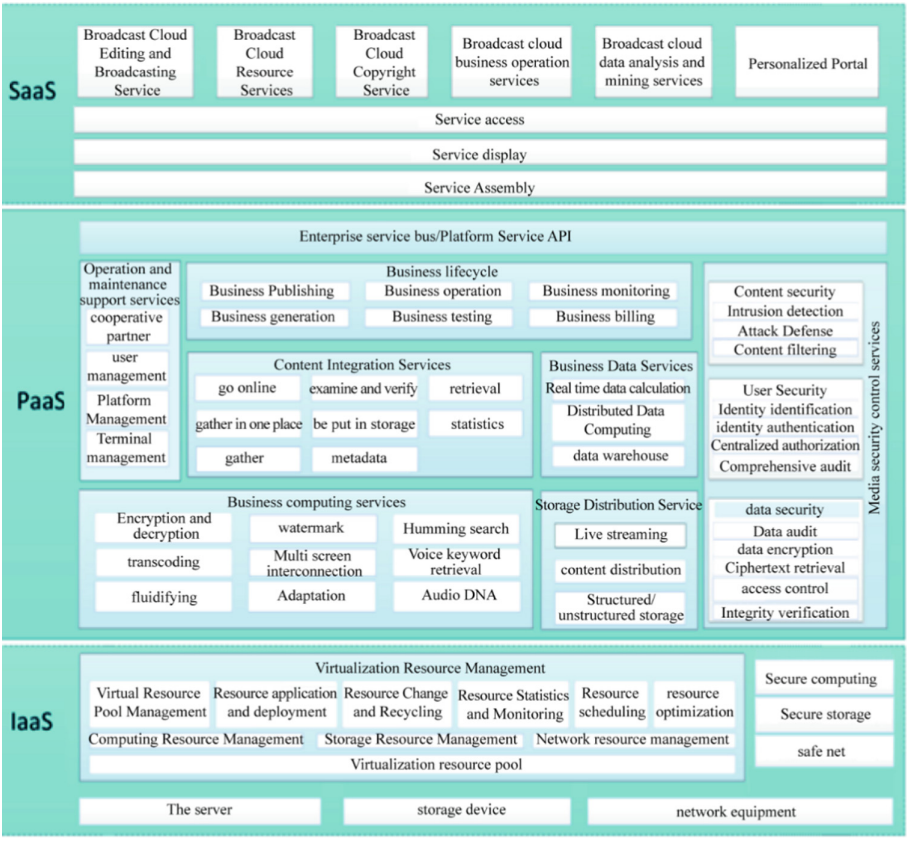

Fig. 2. Overall service architecture design of broadcast private cloud

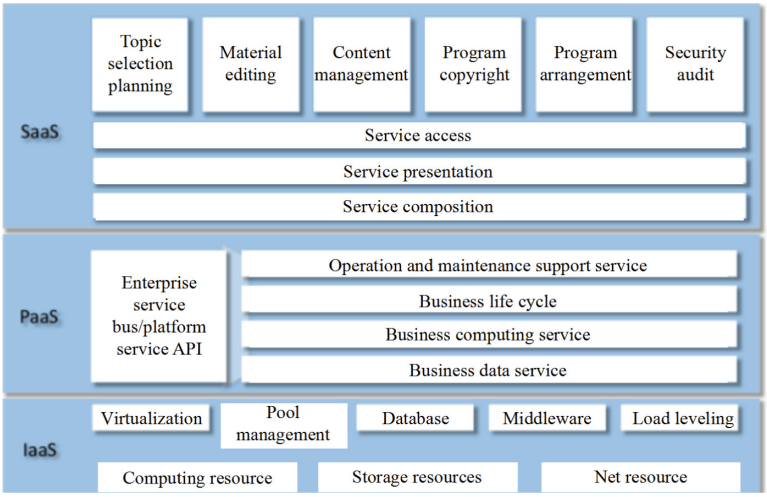

Fig. 3. Schematic diagram of private cloud overall architecture

the DAG scheduler can formally create an execution plan, use the Task scheduler to complete the task assignment, realize the control of the Worker, and formally start executing the parallel computing program.

(5) Data presentation layer: This functional layer can provide richer data query services for end users with the help of Echarts technology, and provide corresponding data analysis results query services. Figure 3 shows the schematic diagram of private cloud overall architecture.

5 Development and Test Analysis of Private Cloud Computing Platform Based on Convergence Media

5.1 Platform Operation Test

In order to ensure the use effect of the cloud computing platform, the working stability of the private cloud computing platform based on OpenStack is tested, and the supporting effect of storing the swift service function is tested. By imitating the service function and storage logic of the enterprise public cloud disk, a new private cloud network disk is specially designed, including the functions of adding and deleting system files and system directories, checking and modifying, uploading and downloading. Finally, the system design is developed and tested through javaweb and the final test results are given. The mainstream open source development framework supported by PaaS is shown in the Fig. 4.

5.2 Create a Database File

(1) The timeliness of information transmission is poor. The rapid development of the economic field has injected a steady stream of development momentum into other

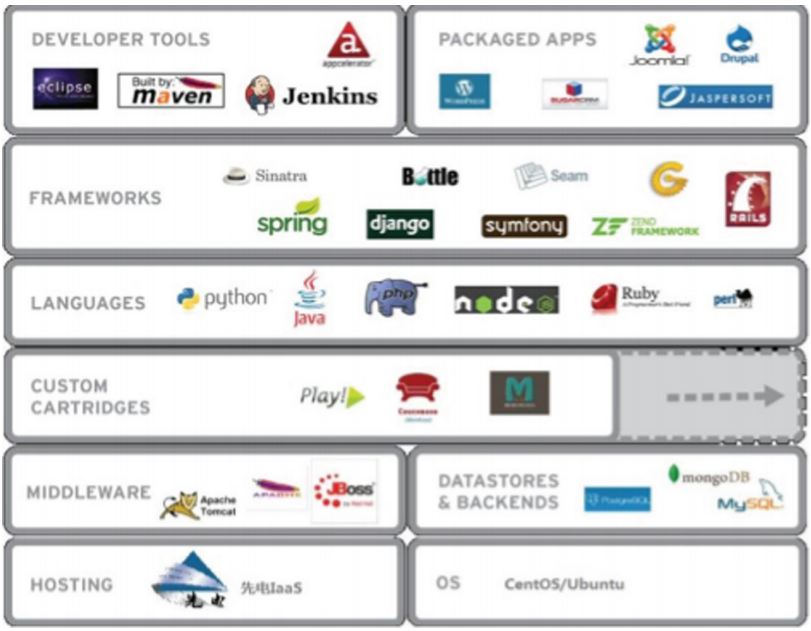

Fig. 4. The mainstream open source development framework supported by PaaS

fields in China, and Internet technology and its related applications have therefore entered a stage of rapid development. In China, the actual coverage of the existing mobile Internet is increasing year by year, and the communication technology in the mobile Internet system has quietly entered the 5G era, and people's dependence on wireless communication technology in their work and life is becoming stronger and stronger. This also represents that the communication technology of modern mobile Internet system (that is, 5G technology) needs more and more electric energy, and the amount of Internet data has changed dramatically. Under such circumstances, the difficulty of information processing in the system is also increasing [6].

(2) The system stability is insufficient. In the Internet of Things system, the sending and receiving of system data is a key content, and it is necessary to ensure the accuracy of data. Therefore, the receiving and sending of data operation instructions can be completed under the control of their respective clock devices (TCLK and RCLK). However, both receiving and sending operations need to be controlled at the same baud rate as the number of characters, and no error can occur. However, the existing information system of Internet of Things has not kept up with the development progress of communication technology in time in data transmission, the application of 5G technology is not complete, and the accuracy of information positioning is insufficient [7].

(3) The establishment of cloud computing platform also needs to build the Internet of Things, which can be used to complete the effective connection between "things and things". Figure 5 shows the schematic diagram of data architecture of cloud computing platform.

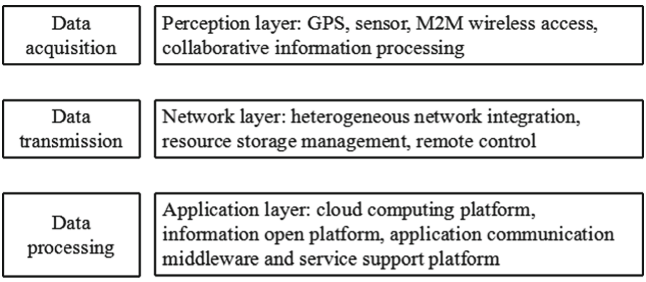

Fig. 5. Schematic diagram of data architecture of cloud computing platform

Perceptual control layer	Sensor→Cable network→GPS IC card→Wireless base station→RFID
Network transport layer	VPN→Local area network→Internet Optical fiber ring network: Primary substation, Secondary substation
Cloud computing service layer	Software maintenance: ERP, CRM, Website, E-commerce Application service: Storage application, Software application, Collaborative application

Fig. 6. Schematic diagram of cloud computing control center with Internet of Things as the core

Designing a more standardized Internet of Things architecture is the core concept of the development and construction of Internet of Things informatization. Based on the traditional hierarchical design concept, a "cloud-network-end" mode setting method can be implemented. The details of the design of Internet of Things informatization architecture combined with big data are shown in Fig. 6:

This layer is mainly composed of HDFS, HBase and Hive. Among them, the underlying storage of data needs to use Hdfs distributed file system. For Hdfs, NameNode should be selected as the designated datanode server and block(64M) according to the actual data size, and all data generated here should be stored in the corresponding block system in turn. In the above process, different types of small files need to be successfully merged into a single large file and then stored. At this time, Hadoop can effectively complete the file merging task with its own CombineFileInput-Format function, and the corresponding data warehousing system needs to be realized by the joint action of Hive and Hbase, and the data analysis results are more targeted [8].

5.3 Data Perception Design

Based on the inherent Internet of Things architecture, expanding and integrating big data technology to form a new information system architecture can deal with all existing problems of the Internet of Things more effectively, and at the same time build a more secure and reliable Internet of Things system with modern, digital and information characteristics, especially after successfully integrating big data technology, improving data management efficiency. The core task of the perception control layer is to effectively collect the information data of multi-dimensional perception nodes, because nodes

belong to the basic organizational form in the Internet of Things system, which is not limited to a front-end perception device, but also can be used as a system front-end with partition function. During this period, because the data resources of the Internet of Things are complex and diverse, the collection tasks and transmission periods of structured data and semi-structured data of the system are different to some extent, and there are significant differences in structure, data collection period, network transmission speed and transmission period in the face of different types of data [9]. Therefore, edge computing terminal equipment can be selected as the transition link between sensor and cloud computing service platform, and the computing power of edge computing itself is outstanding, which can effectively complete the data collection tasks required by the Internet of Things.

In the above content, the network transport layer is between the perception control layer and the cloud computing service layer, which belongs to the transitional link, and virtualization technology can be used to uniformly process all data content [10]. First of all, after completing the data receiving task, the received data needs to be specially processed, and then all the processed data are transmitted to the corresponding cloud computing service layer, which can more effectively improve the data calculation amount provided by the cloud computing service layer and ensure the working efficiency of the cloud computing service layer.

6 Conclusion

To sum up, the successful application of big data technology to the construction of convergence media platform can greatly improve the overall informatization level of existing convergence media. However, because the cloud computing service layer needs too many services, it is not only responsible for the multidimensional data fusion of Internet of Things information data, but also needs to provide services for other functional applications, so the cloud computing task is too heavy. In the above content, the cloud computing service layer will also collect the data transmitted from the network layer to the Internet and the data of service business, which can be regarded as a huge cloud data resource library. Therefore, it is necessary to use big data technology to successfully complete the data analysis tasks proposed by the cloud computing service layer. Finally, it is necessary to rely on the powerful cloud computing service platform and cooperate with big data analysis technology to make a more in-depth and comprehensive analysis and use of the existing network data resources. After in-depth analysis of the internal relationship between data, it can provide various types of data for user terminals, which can be reflected in specific application services, and complete the fundamental task of information construction of convergence media.

References

1. Zhou Yanlin. Research on the Reform of Newspaper Content Production and Public Opinion Guidance Mechanism in the Media-melting Era— Comment on Media-melting Construction and Innovation[J]. China Measurement & Testing Technology,2021,07(03):165-166.

2. Cui Jindong, Li Chenyu, Li Feifei. Research on Hot Spot Discovery Mechanism of Mainstream Media in the Context of Big Data[J]. Information Science,2021,39(12):118-119.
3. Liu Xiaoming. Innovative Application of New Media Technology in the Communication of Ideological and Political Education— Comment on "Innovative Research on Communication of Ideological and Political Education in the New Media Era"[J]. Journalism Lover,2021,05(10):142-143.
4. Chen Jiyin, Cui Yike. Resource Integration and Streamlining: The Construction of County-level Financial Media in Western China under the Multi-synergy—A Study Based on the Construction of County-level Financial Media in X Urban Area[J]. Shanghai Journalism Review,2021,04(12):129-130.
5. Ding Baiquan. Insist on All-round Innovation, Follow Objective Laws and Optimize Content Production — Investigation Report on the Construction and Development of County-level Media Center[J]. Journalism Lover,2022,03(07):137-138.
6. Liu Yuetan, Shi Siqi. "Big Data+Full Live Broadcasting": The Application of Media Convergence Technology in Public Emergency Reporting—Taking "Yongpai+City Brain 74-hour Live Broadcasting to Defend against Typhoon" as an Example[J]. The Press,2022,01(02):154-155.
7. Zeng Xiong. Transformation from Urban Media to Urban Platform—Exploration and Practice of Deep Integration of Changsha Radio and Television Media[J]. The Press,2022,08(03):124-125.
8. Li Zengjun, Liu Yong, Li Xin. Improving the Public Service Level of County-level Media Integration Center—The Practical Exploration of Yumen to Build a Comprehensive County-level Media Integration Demonstration Center in Northwest China[J]. The Press,2021,05(02):163-164.
9. Zhang Zihan, Xie Guangmu. The Combination, Integration and Deep Integration of Information Technology and Chinese Curriculum—Based on the Vertical Comparison of Chinese Curriculum Standards (syllabus) texts[J]. China Educational Technology, 2021,07(08):108-109.
10. Gao Jing. Exploring New Management Methods in Media Technology Development Projects—Taking the Practice of Sichuan Daily's All-media "Smart Party Building" Project as an Example[J]. The Press,2022,05(02):124-125.