International Journal of Management Science Research, Vol.8, Issue 8, (Aug) 2025
ISSN 2536-605X
DOI: 10.53469/ijomsr.2025.08(08).07

Design of Communication Network Data Processing and Analysis Platform Based on Cloud Computing

Tao Yan, Fengbo Tao

Shandong Yellow River Conservancy Bureau East Asia Lake Management Bureau Shandong Tai'an 271000

Abstract: ***With the rapid development of information technology, communication network data is growing explosively. How to efficiently process and analyze this data has become a major challenge facing the industry. This article aims to explore the design of a cloud computing based communication network data processing and analysis platform. By integrating the elastic scalability, on-demand payment, and resource sharing characteristics of cloud computing, an efficient, secure, and reliable data processing and analysis platform is constructed. The article first outlines the basic concepts and interrelationships between cloud computing and big data, and then elaborates in detail on the theoretical basis of platform design, requirement analysis, architecture design, key technology implementation, and information security assurance.***

Keywords: Cloud computing; Communication network; Data processing and analysis platform; Design.

1. INTRODUCTION

In the digital age, communication network data has become an indispensable strategic resource for various industries. Traditional data processing methods are no longer able to meet the demands of massive, high growth, and diverse data. Cloud computing, as an emerging information technology, provides new solutions for communication network data processing with its unique advantages. A cloud computing based communication network data processing and analysis platform can not only significantly improve data processing efficiency, but also reduce costs and enhance data security. This study is of great significance for promoting the intelligent development of communication networks and facilitating the effective utilization of data resources. Tian et al. (2025) pioneered a business intelligence approach using cross-attention multi-task learning to enhance ad recall in digital advertising[1]. In medical imaging, Chen et al. (2023) developed a generative text-guided 3D vision-language pretraining framework for unified medical image segmentation[2]. Recruitment technology was advanced by Li et al. (2025) through their integration of generative pretrained transformers with hierarchical graph neural networks for optimized resume-job matching[3]. Security in IoT systems was addressed by Miao et al. (2025) who designed a secure and efficient authentication protocol for AI-based supply chain systems[4]. Peng et al. (2025) contributed to 3D vision-language understanding through Gaussian Splatting techniques[5]. Financial risk management saw substantial innovations with Su et al. (2025) developing a WaveLST-Trans model for anomaly detection and risk early warning in financial time series[6], while Wang et al. (2025) conducted empirical studies on AI-enhanced intelligent financial risk control systems for multinational supply chains[7,8]. Neural network optimization was advanced by Wu et al. (2023) through Jump-GRS, a multi-phase structured pruning approach for neural decoding[9]. Legal text processing was improved by Xie et al. (2024) with their Conv1D-based approach for multi-class classification of legal citation texts[10]. Network analytics was enhanced by Zhang, Yujun et al. (2025) through MamNet, a novel hybrid model for time-series forecasting and frequency pattern analysis[11]. Green finance applications were addressed by Zhang, Zongzhen et al. (2025) through deep learning approaches for carbon market price forecasting and risk evaluation[12]. Computer vision saw innovations with Zheng et al. (2025) developing Diffmesh, a motion-aware diffusion framework for human mesh recovery from videos[13]. Finally, agricultural technology was advanced by Zhou (2025) through a swarm intelligence-based multi-UAV system for precision pesticide spraying in irregular farmlands[14].

2. THEORETICAL BASIS

2.1 Overview of Cloud Computing and Big Data

Cloud computing is an Internet based computing model. It encapsulates computing resources, storage resources and network resources into an independent virtual environment through virtualization technology to provide users with pay as you go services. This mode allows users to dynamically acquire and release resources based on actual needs, greatly improving the efficiency and flexibility of resource utilization. Big data refers to a collection of data that cannot be captured, managed, and processed using conventional software tools within a certain period of time. It has the characteristics of massive, high growth rate, and diversity. Cloud computing and big data complement each other, providing powerful computing power and storage space for big data processing, making large-scale data processing and analysis feasible and efficient.

2.2 Key Technical Support

2.2.1 Data Mining:

Data mining is the process of extracting useful information from massive amounts of data, discovering correlations and patterns between data. It covers various technologies such as classification, clustering, association rule mining, etc., aiming to help users extract valuable knowledge and insights from complex datasets.

2.2.2 Machine Learning

Machine learning is an artificial intelligence technology that achieves prediction and classification of unknown data through learning and training of data. Machine learning algorithms can automatically extract features from large amounts of data and build predictive models to accurately predict and classify new data.

2.2.3 Distributed Computing

Distributed computing is a computing model that decomposes computing tasks into multiple subtasks and executes them in parallel on multiple computing nodes. It can significantly improve computational efficiency and shorten computation time by utilizing the computing power of multiple computing nodes.

2.2.4 Virtualization Technology

Virtualization technology is a technology that enables dynamic allocation and management of computing resources, storage resources, and network resources. It abstracts physical resources and transforms them into logical resources, thereby achieving flexible management and efficient utilization of resources. Virtualization technology enables cloud computing platforms to dynamically allocate and release resources based on user needs, improving resource utilization efficiency and service flexibility.

3. REQUIREMENT ANALYSIS

3.1 Storage Requirements

Communication network data has the characteristics of massive and diverse. Therefore, the platform needs to have high scalability and elastic storage capabilities to meet the growing demand for data storage [1]. This means that the platform should be able to dynamically expand storage space to cope with sudden increases in data volume and ensure data persistence and availability.

3.2 Processing Capability

Efficient data processing capability is the core requirement of the platform. The distributed computing characteristics of cloud computing can provide users with powerful computing capabilities and enable rapid processing of big data. The platform should fully utilize the distributed computing resources of cloud computing, improve data processing speed through parallel processing and optimization algorithms, and ensure that users can obtain processing results within a reasonable time.

3.3 Security and Privacy Protection

Data security and privacy protection are important aspects that cannot be ignored in platform design. The platform should take various measures to ensure the security of data during transmission and storage, including data encryption, access control, and data anonymization.

3.4 Scalability and Flexibility

The platform should have good scalability and flexibility to adapt to the needs of different business scenarios. This means that the platform should be able to support the integration of multiple data sources and formats, so that users can integrate data from different sources and formats into the platform for unified processing and analysis. At the same time, the platform should also provide rich data processing and analysis functions to meet the diverse needs of users.

3.5 User Interface and Usability

For non professional users, the platform should provide a concise and clear operating interface and comprehensive help documents to lower the threshold for use. The operating interface should be designed to be intuitive and user-friendly, allowing users to easily get started and quickly complete data processing and analysis tasks. At the same time, the platform should also provide detailed help documents and online support so that users can receive timely assistance and solutions when encountering problems.

4. ARCHITECTURE DESIGN

4.1 Hardware Architecture

For the hardware architecture of the platform, it is recommended to adopt a distributed cloud computing architecture. The core advantage of this architecture is that it consists of a large number of computing nodes and storage nodes, which are connected through high-speed networks to form a massive pool of computing resources. This design enables the platform to dynamically allocate and release computing resources based on actual needs, thus flexibly meeting users' data processing and analysis needs in different scenarios. In terms of computing nodes, it is recommended that the platform adopt high-performance server clusters. These servers should have strong computing power and stable operational performance to ensure the smooth progress of data processing and analysis tasks, while improving overall computing efficiency. In order to further improve computational efficiency, parallel computing technology can be considered to enable multiple computing nodes to process the same task simultaneously, thereby significantly reducing the time for data processing and analysis. In terms of storage nodes, it is recommended that the platform adopt a distributed storage system. This system can store data in multiple storage nodes, thereby improving the reliability and availability of data. Even if a storage node fails, other nodes can still function normally, ensuring data integrity and security. In addition, it is recommended that the platform design data backup and disaster recovery mechanisms. This mechanism is crucial for dealing with unforeseeable situations such as hardware failures or natural disasters, ensuring the integrity and non loss of data. It is recommended to adopt a multiple backup strategy, backup data to multiple different storage nodes, and set up a disaster recovery backup center to effectively respond to possible catastrophic events.

4.2 Software Architecture

For the software architecture of the platform, it is recommended to adopt microservice architecture and containerization technology to improve the flexibility and scalability of the system. By splitting the platform into multiple independent services, each responsible for a specific function or business process, decoupling and modularization of services can be achieved. In this way, when specific services need to be added, modified, or replaced, it will not have a significant impact on the entire system, thereby improving the maintainability and scalability of the system. Meanwhile, it is recommended to further adopt containerization technology to package and manage these microservices. Containerization technology can package each service and its dependencies into an independent, runnable container that can run consistently in any environment that supports container operation. This can not only improve the portability and deployability of services, but also achieve isolation between services and effective management of resources. For the convenience of users accessing and analyzing data, it is recommended to design a unified data access interface. These interfaces should provide standardized data access methods so that users can easily access the required data through simple calls [2]. In addition, some commonly used data processing frameworks and algorithm libraries can be provided, allowing users to achieve complex data processing and analysis tasks through simple configuration.

4.3 Network Architecture

For the network architecture of the platform, it is recommended to adopt a flat network topology structure. This structure can reduce network hierarchy, lower network latency and complexity, and improve data transmission efficiency. By reducing routing jumps and network congestion, a flattened network topology can ensure fast data transmission in the network, thereby improving the overall performance of the platform. To ensure the security of data transmission, it is recommended to use encryption technology in the network architecture. By encrypting the transmitted data, it can prevent data from being stolen or tampered with during transmission, protecting users' privacy and sensitive information. The application of encryption technology can effectively improve the security of data transmission and enhance users' trust in the platform. Meanwhile, to ensure the stability of data transmission, it is recommended to introduce load balancing technology into the network architecture. Load balancing technology can evenly distribute network traffic to multiple servers, avoiding network congestion and service unavailability caused by overloading a single server. Through load balancing, the stability and reliability of data transmission can be ensured, and the availability and user experience of the platform can be improved.

5. KEY TECHNOLOGY IMPLEMENTATION

5.1 Cloud computing infrastructure

Cloud computing infrastructure is the core of utilizing virtualization technology to dynamically allocate and manage computing resources, storage resources, and network resources. Through virtualization technology, physical servers, storage devices, and network resources are abstracted into logical resource pools that can be dynamically allocated and adjusted according to demand. Automated deployment tools are used to quickly build and scale cloud computing environments, improving deployment efficiency and resource utilization. Specifically, it includes:

(1) Virtualization technology:

Use virtualization software such as VMware and Hyper-Vto partition physical servers into multiple virtual machines, each of which can independently run operating systems and applications. Meanwhile, utilizing virtual storage and virtual network technology to achieve virtualization of storage and network resources.

(2) Automated deployment tool:

Using automated configuration management tools such as Ansible and Puppet, as well as continuous integration/continuous deployment (CI/CD) tools such as Jenkins and Kubernetes, to achieve rapid construction, configuration, and scaling of cloud computing environments.

5.2 Big Data Storage and Management

Big data storage and management are key to platform processing of large-scale data. By adopting distributed file systems (such as Hadoop HDFS) and NoSQL database technologies, efficient data storage and access can be achieved. At the same time, the concept of data lake is introduced to store data from different sources and formats uniformly in the data lake, which facilitates subsequent data processing and analysis. Specifically, it includes:

(1) Distributed File System:

Using Hadoop HDFS and other distributed file systems to store data on multiple nodes in a decentralized manner, achieving parallel access and processing of data. HDFS has the characteristics of high fault tolerance and high throughput, making it suitable for large-scale data storage and processing [3].

(2) NoSQL database technology:

Using NoSQL databases such as MongoDB and Cassandra to store non relational data. NoSQL databases have high performance and scalability, making them suitable for fast reading and writing of large-scale data.

(3) Data Lake:

Build a data lake to store data from different sources and formats in a unified manner. A data lake can provide a unified data access interface, facilitating subsequent data processing and analysis. At the same time, the data lake also supports real-time updates and queries of data, meeting the needs of real-time data processing.

5.3 Data Processing and Analysis

Data processing and analysis are the core functions of the platform. Develop efficient data processing and analysis algorithms by combining data mining and machine learning techniques. Utilizing the distributed computing characteristics of cloud computing to accelerate data processing and analysis processes. Specifically, it includes:

(1) Data mining techniques:

Using data mining techniques such as association rule mining and cluster analysis to extract valuable information and patterns from large-scale data. These technologies can help users discover potential patterns and correlations in data, providing support for decision-making.

(2) Machine learning technology:

Train and predict data using machine learning algorithms such as decision trees and neural networks. Machine learning technology can automatically learn and optimize data processing and analysis processes, improving the accuracy and efficiency of data processing.

(3) Distributed computing:

By utilizing the distributed computing characteristics of cloud computing, data processing and analysis tasks are assigned to multiple computing nodes for parallel execution. This can greatly improve the speed and efficiency of data processing and analysis, meeting the needs of real-time data processing and analysis. Meanwhile, distributed computing can also improve the scalability and fault tolerance of the system, ensuring the stable operation of the platform.

6. SECURITY MEASURES

6.1 Data Encryption

To ensure the security of data during transmission and storage, advanced encryption algorithms can be used to encrypt the data. For data in transmission, TLS/SSL protocol is used for encryption to ensure that the data is not stolen or tampered with during transmission. For stored data, strong encryption algorithms such as AES-256 are used to ensure that even if the data is illegally obtained, it cannot be easily decrypted.

6.2 Access Control

To strictly restrict access to sensitive data, strict access control policies can be implemented. Adopting the Role Based Access Control (RBAC) model, users are divided into different roles and assigned different access permissions to each role [4]. Only authorized users can access data within their authorized scope, ensuring that sensitive data is not accessed by unauthorized users.

6.3 Data Desensitization

In order to further reduce the risk of data leakage, it is necessary to desensitize sensitive data. By using techniques such as data replacement and data scrambling, sensitive data can be transformed to prevent direct recognition or restoration without changing the original data structure and patterns. In this way, even if data is leaked, sensitive information cannot be directly obtained.

6.4 Emergency Response

In order to respond to potential information security incidents, detailed emergency plans should be developed. The contingency plan includes emergency response procedures, responsible persons, and recovery measures to ensure prompt and effective response in the event of an information security incident. Establish a dedicated information security emergency response team responsible for handling various information security incidents.

7. CONCLUSION

This article elaborates on the design and research of a cloud computing based communication network data processing and analysis platform. By integrating the elastic scalability, on-demand payment, and resource sharing features of cloud computing, combined with key technologies such as data mining, machine learning, and distributed computing, an efficient, secure, and reliable data processing and analysis platform has been constructed. This platform not only meets the diverse needs of communication network data processing, but also significantly improves data processing efficiency and security, providing strong support for the intelligent development of communication networks. In the future, with the continuous advancement of technology and the expansion of application scenarios, cloud computing based data processing and analysis platforms will play a more important role.

REFERENCES

[1] Q. Tian, D. Zou, Y. Han and X. Li, "A Business Intelligence Innovative Approach to Ad Recall: Cross-Attention Multi-Task Learning for Digital Advertising," 2025 IEEE 6th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Shenzhen, China, 2025, pp. 1249-1253, doi: 10.1109/AINIT65432.2025.11035473.

[2] Chen, Yinda, et al. "Generative text-guided 3d vision-language pretraining for unified medical image segmentation." arXiv preprint arXiv:2306.04811 (2023).

[3] Li, Huaxu, et al. "Enhancing Intelligent Recruitment With Generative Pretrained Transformer and Hierarchical Graph Neural Networks: Optimizing Resume-Job Matching With Deep Learning and Graph-Based Modeling." Journal of Organizational and End User Computing (JOEUC) 37.1 (2025): 1-24.

[4] Miao, Junfeng, et al. "Secure and Efficient Authentication Protocol for Supply Chain Systems in Artificial Intelligence-based Internet of Things." IEEE Internet of Things Journal (2025).

[5] Peng, Q., Planche, B., Gao, Z., Zheng, M., Choudhuri, A., Chen, T., Chen, C. and Wu, Z., 3D Vision-Language Gaussian Splatting. In The Thirteenth International Conference on Learning Representations.

[6] Su, Tian, et al. "Anomaly Detection and Risk Early Warning System for Financial Time Series Based on the WaveLST-Trans Model." (2025).

[7] Wang, Z., Chew, J. J., Wei, X., Hu, K., Yi, S., & Yi, S. (2025). An Empirical Study on the Design and Optimization of an AI-Enhanced Intelligent Financial Risk Control System in the Context of Multinational Supply Chains. Journal of Theory and Practice in Economics and Management, 2(2), 49–62. Retrieved from https://woodyinternational.com/index.php/jtpem/article/view/208

[8] Wang, Zhiyuan, et al. "An Empirical Study on the Design and Optimization of an AI-Enhanced Intelligent Financial Risk Control System in the Context of Multinational Supply Chains." (2025).

[9] Wu, Xiaomin, et al. "Jump-GRS: a multi-phase approach to structured pruning of neural networks for neural decoding." Journal of neural engineering 20.4 (2023): 046020.

[10] Xie, Y., Li, Z., Yin, Y., Wei, Z., Xu, G., & Luo, Y. (2024). Advancing Legal Citation Text Classification A Conv1D-Based Approach for Multi-Class Classification. Journal of Theory and Practice of Engineering Science, 4(02), 15–22. https://doi.org/10.53469/jtpes.2024.04(02).03

[11] Zhang, Yujun, et al. "MamNet: A Novel Hybrid Model for Time-Series Forecasting and Frequency Pattern Analysis in Network Traffic." arXiv preprint arXiv:2507.00304 (2025).

[12] Zhang, Zongzhen, Qianwei Li, and Runlong Li. "Leveraging Deep Learning for Carbon Market Price Forecasting and Risk Evaluation in Green Finance Under Climate Change." Journal of Organizational and End User Computing (JOEUC) 37.1 (2025): 1-27.

[13] Zheng, Ce, et al. "Diffmesh: A motion-aware diffusion framework for human mesh recovery from videos." 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). IEEE, 2025.

[14] Zhou, Dianyi. "Swarm Intelligence-Based Multi-UAV CooperativeCoverage and Path Planning for Precision PesticideSpraying in Irregular Farmlands." (2025).