Design and Application of Fog Computing Model Based on Big Data

Baoling Qin, Xiaowei Lin, Sina Li, Qiao Luo, Fenglin Zheng, JieJian Cai, Yunshi Luo
College of Electronic Information Engineering, Foshan University
Foshan, China
e-mail: 2978879663@qq.com,1691139513@qq.com,144524241754@qq.com,1075855518@qq.com,
1225288937@qq.com,1150074271@qq.com,1091024184@qq.com

***Abstract*—Fog computing based on big data is a hot topic in the research of computing technology at home and abroad. With the wide application and popularity of IoT (Internet of Things), the big data generated by edge devices is exploding, and cloud computing models are becoming increasingly inadequate to meet the needs of big data processing and communication, which is mainly manifested as follows. Slow data processing, insufficient storage space, prolonged communication and many other issues. Fog computing, of which the advantage is distributed computing, namely the "de-centralized" mode calculation, is the suitable solution to solve these problems. In the IoT system, fog computing model based on big data is constructed to distribute the big data computing, storage and communication in the system to the edge device. The purpose is to make the system structure simpler, more modular and intelligent, duce network congestion, exploit advantages of edge devices and improve high quality intelligence of IoT applications, and moreover, to reduce the deployment of IoT hardware and operating costs. Taking the cloud robotics as an example, it is proposed to embed the fog computing technology in the cloud robotics system, which greatly improves the computing function of the cloud robotics system. In short, it provides theoretical support and scientific experimental basis for the informationization and intelligence of all walks of life, and its research has certain value and significance.**

Keywords-big data; fog computing; cloud computing; internet of things; fog models

I. INTRODUCTION

With the development of the era of big data, the connection of the Internet of Things in people's lives is getting closer and closer. Especially in the context of large amounts of data generated by sensors, cloud computing-based big data processing and communication show obvious shortcomings. For example, slow processing, insufficient storage space, high latency, etc. The solution to these serious problems is to build a distributed structure model of big data processing based on fog computing and fully utilize the advantages and features of its "de-centralization" to solve the problems caused by cloud computing in the process of big data processing, making the structure of the IoT application process clearer, simpler and more reasonable, the information processing process faster, smoother, more reliable, and the fields involved more extensive, which also brings certain scientific basis and development to its field research.

II. CONCEPT AND TECHNICAL ANALYSIS OF BIG DATA FOG CALCULATION

Big data exists everywhere in the data age. As an extension of cloud computing, fog computing is ubiquitous in practical applications. Nowadays, the research and application of big data and fog computing can be said to be enthusiasm and unstoppable. Though there exist differences in understanding and definition of them in the progress and development of the times, they are all similar.

A. Big Data

With the rapid development of the times, big data is increasingly related to people. Big data can be regarded as a data set, which integrates complex and large-scale data structures and existing database managers or data processing applications are difficult to properly deal with. As will be readily seen, big data features large data volume, complex structure, high practicability, strong diversity, etc. It mainly comes from various types of data generated by people in the process of using the Internet, such as pictures, audio, and so on; as well as data generated by mobile phones, computers and other terminal equipment in the form of documents, multimedia, etc.; there are also a number of data generated by hardware devices such as videography, medical care equipment, etc. Its applications cover a wide range of fields, such as medicine, business, science, etc. Making use of big data, we can obtain the corresponding knowledge and predict future trends and analyze and master the personalized characteristics of the field. Besides, utilizing the diversity of big data, we can distinguish the true and false acquired information.

B. Cloud Computing

The wide application of the Internet of Things brings a lot of terminal equipment investment, and a large amount of data comes one after another. In the face of the huge amount of data that needs to be processed, Google engineer Christopher Bihlia proposed the concept of Google "cloud computing" in the "Google 101" project [1]. Through the Internet, cloud computing can provide real-time, fast, convenient, and secure use of shared computing facilities, storage devices, applications and other resources to provide powerful processing data computing and storage capabilities.

978-1-7281-3323-2/19/$31.00 ©2019 IEEE

It provides fast, convenient, real-time, on-demand network services. Cloud computing has the advantages of virtualization, large scale, on-demand service, low price, high scalability, high reliability, etc., which greatly facilitates customers to use cloud computing for data processing and analysis.

With the rapid development and widespread application of the network, the Internet users have exploded, and the number of terminal devices accessing the network has increased. But at the same time, the network congestion has been caused by the squeeze of bandwidth. The data transmitted to the "cloud center" occupies bandwidth, which is not processed through a large number of sensors in the Internet of Things, and of which a large part is not necessary to be transmitted. Many devices and applications require low latency while cloud computing cannot meet the low latency and cannot meet the user's demand for network speed, and the computing capabilities and data of cloud computing are concentrated in the "cloud center". Data is occasionally lost, leaked, etc.

C. Fog Computing

In response to some shortcomings of cloud computing, in 2012, Cisco's F. Bonomi et al. first proposed the concept of fog computing [2]. This idea is not to replace the cloud computing model, but to supplement cloud computing. Because cloud computing is mainly composed of backbone network and data center, fog computing will be further extended to include edge networks composed of various IoT networks, and devices, smart meters and smart sensors connected by these edge networks, etc. on the basis of it. Computing is mainly based on three types of small clouds, such as personal cloud, family cloud and unit cloud, rather than the "big cloud" and "social public cloud" of IT service providers advocated by early cloud computing.

Fog computing can be perceived in huge cloud computing systems and big data structures. It can extract limited and accurate distribution content, improve the accuracy of problem solving, achieve higher real-time data calculation and storage, and greatly lighten computing and storage burden of cloud computing.

1) Characteristics of fog computing technology

a) Low latency. Because it is on the edge, close to the customer, the communication delay is very low, which is of great significance for the booming Internet of Things, suitable for industrial intelligence, online games, video transmission, augmented reality, etc.

b) Dense and extensive geographical distribution. This characteristic makes it possible to achieve more efficient service requests for mobile users, while the cloud computing distribution is sparse.

c) Massive nodes. A large-scale sensor network with a large number of network nodes facilitates monitoring of the environment.

d) Quickly wireless access. For fog computing, terminal devices such as mobile phones and other mobile devices can communicate directly with each other via Bluetooth, WiFi, etc, and the signals do not have to go to the cloud or even the base station to cycle through the circles, which greatly improve transmission efficiency.

e) Real-time analysis and control. Real-time interaction and online cloud analysis are supported.

f) Energy and money saving. The nodes of the fog computing are decentralized and do not require an additional cooling system to handle the heat to maintain balance outside the fog server.

g) Heterogeneity. Different shapes and different environments support a variety of heterogeneous hardware and software devices.

2) Advantages of fog computing technology

a) Confidentiality. The client usually transmits the information to the cloud directly, but if the client transmits the information to the fog end and the information is filtered, calculated, stored at the fog end, and then transmitted to the cloud, the confidentiality of user information can be guaranteed.

b) Reliability. Data is transmitted over long distances between the data and the cloud layer, sometimes with calculation errors and data loss. Pre-calculating the data on the fog side greatly improves the reliability of data transmission.

c) Convenience. End users can communicate directly via wireless connections such as Bluetooth and Wi-Fi.

d) Low latency .A large number of nodes of fog computing are distributed in the user's nearest place, and can receive the user's service request as soon as possible.

From the above, fog computing is not as complex as cloud computing. It is a cloud-based service model, which leverages the virtualized infrastructure close to the end user and processes it in the cloud to reduce costs and reduce latency, and more visits. The fog calculation concentrates on the advantages of cloud computing and improves the shortcomings of cloud computing to fill the gap in cloud computing.

III. Research and Design of Fog Computing Model Based on Big Data

As mentioned above, compared with cloud computing, fog computing has the function of distributed computing, that is, "decentralization". The next part will be combined with the fog computing architecture mentioned by Cisco. Research and design is based on fog computing model based on big data.

Big data technology has developed rapidly in recent years, which plays an irreplaceable role in human social life. It is the starting point and starting point of research and design of fog computing model how to meet the needs of big data processing and information exchange, and how to solve the current situation of big data technology such as slow data processing, insufficient storage space and prolonged communication.

A. Design of Fog Computing Model Based on Big Data

The fog computing structure is multi-layered. The fog computing system is mainly composed of many individual

fog nodes that are structurally heterogeneous. The scattered fog nodes collect data on the device at the edge of the network and analyze it quickly. The operability of the fog nodes is enhanced by abstract protocol, which achieves access and management among heterogeneous platforms. Fog nodes are divided into advanced and low-level nodes and each of them has a management agent API interface. After the interface is running, the advanced nodes can monitor and control the lower-level nodes, so as to grasp the global situation and instantly understand and analyze the operation and service status of components of the hardware resource device, and then establish an independent management system. In addition, there exists the distributed management database in the node management layer, so that the service capability of the fog nodes can be balanced with the platform capability, which can avoid balance between load and idle, and has flexible space for fault tolerance, and evaluates the status of nodes and component services. In this way, the IoT application system structure will be more simple, modular and intelligent.

Fog computing model based on big data is divided into four layers, that is, the Terminal Perception Layer, the Network Transport Layer, the Data Processing Layer, and the Terminal Application Layer, as shown in Fig. 1.

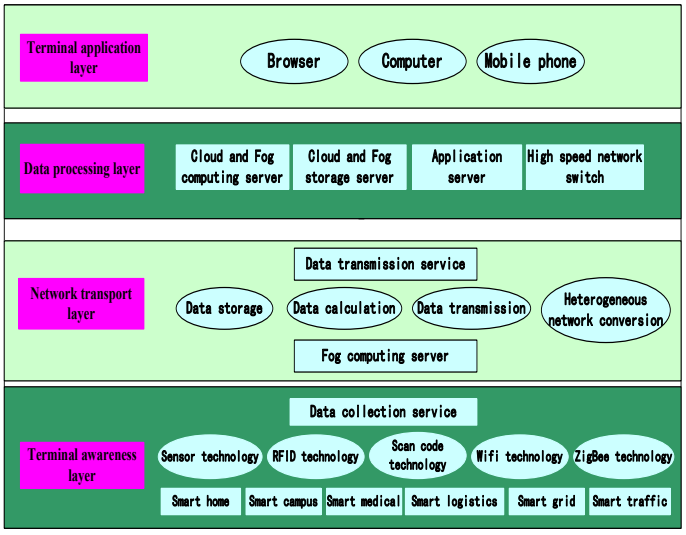

Figure 1." Fog computing model based on big data.

B. Technical Analysis of Each Layer in the Model Structure

There are strong correlations between the layers in the fog computing model structure, and the technical characteristics of each layer are different. The analysis of each layer technology is helpful to better understand the application of the model.

1) Terminal perception layer

The Terminal Perception Layer of fog computing (also known as the Comprehensive Perception Layer) is also the basic source of fog computing data information. The Terminal Perception Layer mainly uses RFID reading, barcode reader, sensor node, intelligent terminal and other technologies for data collection. In addition to sensing data collection, the Comprehensive Perception Layer also includes sensing device networking which can improve the invulnerability and flexibility of the network and enhance the capacity of resisting disturbarice in the transmission process. The main technologies include sensor technology, RFID technology, scanning code technology, ZigBee and Bluetooth technology; the main applications are smart campus, smart grid, smart home, smart logistics, intelligent transportation, smart medical, smart agriculture, smart car networking and so on.

2) Network transport layer

The function of the Network Transport Layer mainly includes the mutual integration of the sensor network and the Internet, and the processing of sensor information [3]. Fog computing is required as a communication bridge between the data and the network. Like cloud computing, fog computing requires the completion of important tasks such as pre-processing, aggregation, storage, analysis, and transmission of data collected by edge devices. In other words, the main functions of the fog computing technology in the Network Transport Layer are as follows: data processing, storage, computing, transmission, heterogeneous network conversion and so on. Applying the fog computing technology to the Network Transport Layer can fully make the IoT application system structure more simple, modular and intelligent.

Specifically, in order to make the data in the fog computing model safe and unimpeded to be sent to the cloud layer, it is first necessary to realize the conversion between the external networks such as the sensor network, the Internet, and the mobile communication network [4]. Under this precondition, only by transforming heterogeneous network can various protocols be supported, such as TCP/IP protocol and ZigBee protocol.

The data collected by the edge device is uploaded to the Fog Computing Layer through the Transport Layer before data storage and data mining. The core lies in the "first hurdle" of fog computing, that is, data preprocessing, including integration of complementary information, deletion of redundant data and intelligent analysis data, which greatly improves the transmission speed and effectively alleviates data congestion and occupation. The phenomenon of large space.

3) Data processing layer

As mentioned earlier, fog computing has preprocessed data as it is transferred to the data center. The Data Processing Layer is the closest to the Cloud Layer in the fog computing structure and can be regarded as a small cloud. It reprocesses and analyzes the massive data of the Network Transport Layer, including cloud storage servers, cloud computing servers, application servers, high-speed network switches and so on. The introduction of the Fog Computing Layer in the Data Processing Layer can solve the problem of massive calculation in cloud computing and reduce the burden of the cloud.

4) Terminal application layer

The Terminal Application Layer is the closest layer to the user in the fog computing model based on big data, and is also the most socially valuable layer. Users transmit the

data requests to the nearest fog nodes through short-range communication, Ad hoc technology, ZigBee technology, etc., and each node manages each other's resources, shares each other's user request work, and processes the received data within a set time interval, and then feeding back to the intelligent terminal user to achieve the effect of filtering data that is not necessarily uploaded to the cloud, reduce the network delay and improve the communication speed. Thus it can be seen that the main functions of the Application Terminal Layer are to control and make decisions on the data of the Data Processing Layer, realize intelligent application and management, and serve it in various platforms and various industries.

C. *Distributed Computing Technology in the Model*

Since the computing power of each edge device in the fog computing model is relatively weak, in the context of big data processing and communication, it is necessary to improve the comprehensive computational capability of its model, and introduce distributed computing technology. This technical feature: Migrating all parts of computing tasks to edge devices for calculation, reducing the heavy computing tasks of the cloud servers, saving the overall computing time and improving the computing efficiency. Besides, the storage space resources of cloud servers are limited and the information is pre-processed and stored locally on the edge devices to reduce the pressure of cloud server storage. In addition, because the calculation and storage of information are carried out in the localized edge device, the information processing service delay rate is greatly reduced, effectively solving the network congestion problem and improving the communication capability.

IV. Application Example Analysis

Fog computing model based on big data has a wide application field, especially for industrial intelligence. The future application of 5G will also refer to the fog computing model based on big data.

A. *Industrial Robot Application Based on Fog Computing*

Industrial robots are multi-joint manipulators or multi-degree-of-freedom machines for industrial applications. Performing work automatically, they, can be commanded by humans or run in accordance with pre-programmed procedures, are machines that realize various functions by their own power and control ability. Modern industrial robots can also act on the principles of artificial intelligence.

Intelligent machines have opened up the era of fog computing applications. Fog computing refers to the storage, transmission and calculation between the terminal devices connected to the Internet of Things and the cloud computing that stores data. With the advancement of the Internet of Things, sensors have become quite sophisticated, and they can now collect large amounts of data. Using fog computing or fog analysis, the intelligent machine performs a part of the analysis locally and delivers the analysis results to the cloud in time. In short, from 2016, fog computing is gaining momentum in smart machine applications [5].

At the Humanoids2010 conference, Professor James Kuffner of Carnegie Mellon University proposed the concept of "cloud robotics" and aroused wide discussion. In other word, the cloud robotics is "cloud + robot". The main purpose of the cloud robotics is to offload highly complex calculations to the cloud platform through network communication technology, thus greatly reducing the computational load of a single robot. Fig. 2 depicts the architecture diagram of the cloud robotics system.

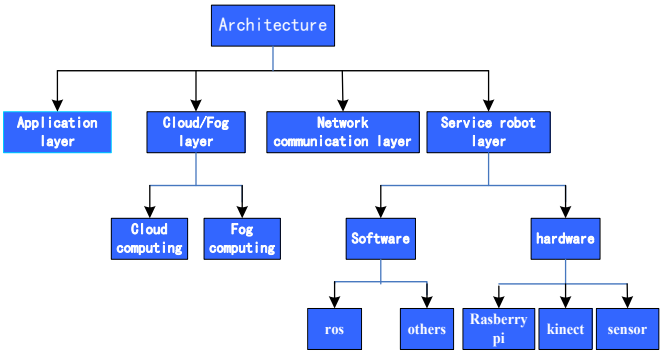

Figure 2." Architecture diagram of the cloud robot system.

The system mainly includes four layers from the bottom up (as shown in Fig.1), the Service Robot Layer, the Network Communication Layer, the Cloud Platform Layer and the Application Layer, wherein each layer communicates through the Communication Network Layer. The Cloud Platform Layer plays a particularly important role in the whole system. The cloud and fog platform provides a large amount of data information for the cloud robotics, and all calculations, fusion analysis, behavior control, resource scheduling, etc. of the cloud robotics are executed in the cloud.

The cloud robotics system architecture involves multiple technologies such as cloud, network and embedded systems, as well as various wireless communication switching management technologies. Here, we mainly discuss two major driving forces: fog computing and big data.

Based on the deficiencies of the cloud computing environment, fog computing is introduced to improve the interactivity and real-time performance of the system. Fog computing is a highly virtualized platform that provides computing, storage, and networking services between end devices and traditional cloud computing data centers, typically but not entirely at the edge of the network [6]. "Fog" is often seen as a "cloud" close to the ground. Similarly, fog computing can also be seen as a cloud computing model in which cloud data centers are close to terminal devices such as robots. According to the literature, a fog computing architecture is shown in Fig. 3.

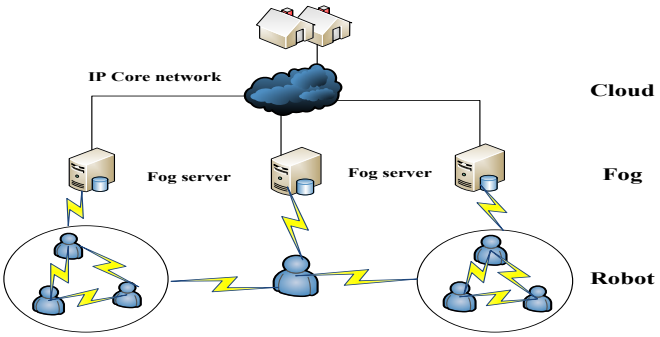

Figure 3." Fog calculation architecture.

The architecture adopts a two-tier computing model of remote cloud and local cloudlet. Compared with the traditional centralized cloud computing architecture, fog computing architecture has obvious localized distributed computing features. Cloud robots use the characteristics of fog computing models based on big data. By offloading computationally intensive tasks to the fog platforms, robots no longer need complex on-board service center equipment. The application cost of multi-robot systems is greatly reduced. It is believed that there will be "Fog Robots", that is, "Fog Computing + Robots".

B. Other Field Applications

For example, in March of last year, media reported that Newifi New Routing launched an intelligent router based on the fog computing model, enabling the network bandwidth to be extended to T-level remote data device layout. Users who purchase the router and connect to a mobile hard disk of 200G or more can use the "Gem Mine" function to experience the "Idle Bandwidth Conversion Network". Similarly in April of last year, the Internet of Things Innovation Wulian Company developed a large-scale intelligent management solution, which was customized by the Tailong Group. Based on the unique "fog computing", the program first introduced the "fog computing host" to build a huge intelligent network and control the network system devices. For example, in November of last year, Yibin County Committee of Sichuan Province explored a new way to educate and train for Party members and cadres, and constructed the first "Electronic Network Ladder" platform for Party members. The platform adopts the education model of "fog calculation + ecological closed loop of Party building", which is comparable with the traditional education informationization The learning mode platform was developed and put into use by Shenzhen Pilot Online Technology Company. It can be said that 5G is an important application milestone of fog computing this year. The solution to incorporate fog computing into the wireless access network (F-RAN) architecture is to use F-RAN as the 5G wireless access network. In short, the era of application of fog computing based on big data has come, and the future is vast.

V. CONCLUSION

By introducing the concepts and development of big data, cloud computing, and fog computing, we can better understand the technical characteristics and technical advantages of fog computing. By analyzing and applying these advantages, a fog computing model based on big data is built to alleviate the pressure of cloud computing on computing, storage and communication in big data processing and information exchange. The purpose is to make the structure of the IoT application system simpler, more modular and intelligent, so that the edge devices of each application can better exert their advantages and improve the intelligence of information application in various industries. At the same time, it can reduce the deployment of IoT hardware and reduce operating costs, provide theoretical basis support and scientific experimental basis for application intelligence and informationization in various industries. Its research results have certain value and significance.

In terms of application, the main analysis is based on the cloud robot example. It is proposed to embed the fog computing technology in the cloud robot structural system, which greatly improves the computing function of the cloud robot system and embodies the important role of fog computing in the cloud robot system. With the development of fog computing, big data, etc., these technologies have spread all over the industrial, commercial and other fields, contributing to the construction of intelligent industries.

REFERENCES

[1]" Chen K, Zheng W M. Cloud Computing: System Instances and Current Research [J]. Journal of Software, 2009,20(5):1337-1348.

[2]" Bonomi F, Milito R,zhu J,et al.Fog Computing and its Role in the Internet of Things.Edition of the Mcc Workshop on Mobile Cloud Computing. ACM, 2012:13-16.

[3]" Yin Yanlin. Research and implementation of embedded intelligent gateway based on Internet of Things [D]. Beijing University of Technology, 2013.

[4]" Shi Yingjuan. Software Architecture Analysis and Application Research of Fog Computing Platform [D]. Zhejiang Normal University, 2016.

[5]" 2016 Big Data Development Trend_The Age of Fog Calculation Analysis Opens[J]. Journal of Distance Education, 2016(1): 60-60.

[6]" Bonomi F,Milito R,Zhu J,et al.Fog computing and its role in the Internet of things[C]//Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, 2012:13-16.