

Software Architecture of Fogcloud Computing for Big Data in Ubiquitous Cyberspace

Jia Yan¹, Fang Binxing², Wang Xiang¹, Wang Yongheng³, An JingBin³, Li Aiping¹, Zhou Bin¹

1. National University of Defense Technology, Changsha 410073

2. Chinese Academy of Engineering, Beijing 100088

3. Zhejiang Lab, Hangzhou, 310012

Abstract: Cyberspace has expanded from the traditional Internet to a ubiquitous cyberspace that connects human, machines, things, services, and applications. The computing paradigm is also shifting from centralized computing in the cloud to combined computing in the front end, middle layer, and cloud. Therefore, traditional computing paradigms such as cloud computing and edge computing can no longer satisfy the evolving computing needs of big data in ubiquitous cyberspace. This paper presents a computing architecture called fogcloud computing for big data applications in ubiquitous cyberspace. Collaborative computing by multiple knowledge actors in the fog, middle, and cloud layer can be realized based on collaborative computing language and models, thereby providing better solutions for big data computing in ubiquitous cyberspace.

Keywords: fogcloud computing; ubiquitous cyberspace; big data; Internet of Things; cloud computing

1 Introduction

Information technology is a catalyst that promotes the development of industry. With the development of cyberspace from the information-oriented Internet to ubiquitous cyberspace, traditional computing architectures like high-performance computing [1], cloud computing [2], fog computing [3], edge computing [4], and multi-agent systems [5,6] cannot meet the needs of big data analysis and processing in cyberspace.

1.1 The expansion of cyberspace

With the rapid development of the Internet, the mobile Internet, the Internet of things, and social networks, cyberspace is becoming more and more prosperous. This promotes the development of cyberspace from the Internet into ubiquitous cyberspace. Ubiquitous cyberspace is an adaptive intelligent network. It connects with the Internet of things, the Internet, sensor networks, and so on through various wired and wireless networks. It comprehensively utilizes terminals such as massive sensors and intelligent processing equipment, as well as the softwares, services, and applications on these terminals, to safely and effectively connect human beings, machines, things, and information at any time and place. Owing to the expansion of the network space, traditional cloud

Received date: August 31, 2018; **Revised date:** August 15, 2019

Corresponding author: Wang Xiang, Research Assistant of the National University of Defense Technology. Major research field is data mining; E-mail: xiangwangcn@nudt.edu.cn

Funding program: CAE Advisory Project (2016-ZCQ-10); Key Project of the National Natural Science Foundation of China “Research on Heterogeneous Media Search Theory and Method for Complex Query” (61732004); General Program of the National Natural Science Foundation of China “Research on Key Technologies of Data Security and Privacy Protection for Cloud Computing Platform” (61472433); Key Project of the National Natural Science Foundation of China “Basic Theory and Technology of Ubiquitous Network Safe Search” (61732022)

Chinese version: Strategic Study of CAE 2019, 21 (6): 114–119

Cited item: Jia Yan et al. Software Architecture of Fogcloud Computing for Big Data in Ubiquitous Cyberspace. *Strategic Study of CAE*, <https://doi.org/10.15302/J-SSCAE-2019.10.001>

computing performed in a background center cannot meet the needs of ubiquitous computing, which contains the Internet, the Internet of things, mobile Internet, and so on. The computing paradigm must change from background processing in cloud centers to a mixture of computing in fog side, middle layer, and cloud centers.

1.2 The characteristics of big data

The term “big data” [7] refers to data that cannot be processed or analyzed using traditional processing methods and tools. We analyze the challenges in the computing of big data in ubiquitous cyberspace by the five characteristics of big data:

(1) *Volume*. The volume of data increases from terabytes to petabytes or even zettabytes. It is difficult for a single application or data center to process such a large amount of data in a timely manner, so it is urgent to realize the collaborative processing of multi-knowledge agents.

(2) *Variety*. There are a larger number of different formats of data, such as documents, emails, social media text messages, videos, still images, audio, graphs, as well as machine-generated data from sensors, devices, RFID tags, machine logs, cell phone GPS signals, and DNA analysis devices. This requires the fusion processing of numerous different data formats. This is too hard for a single person, company, or organization to process. It requires the collaboration of many persons, companies, and organizations while protecting their intellectual property.

(3) *Velocity*. In big data applications in cyberspace, data from human beings, machines, and objects are generated at high speeds, and sometimes must be processed in real time. This makes it important to process data in the fog layer or middle layer which are not far from the source, and maintain knowledge and information interaction with the cloud.

(4) *Veracity*. Data in cyberspace can be uncertain and untrustworthy, so it is vital to perform trusted computing and cross-verifications of different kinds of data.

(5) *Value*. There is considerable value in big data, but it is hard to mine results from a large dataset. Data in cyberspace is related to numerous different domains, making it hard for a single person, company, or organization to mine its value.

In brief, handling big data in cyberspace requires many persons, companies, and organizations to work together while protecting their intellectual property and the ownership of data. Existing architectures and computing paradigms cannot meet the needs of big data computing in cyberspace.

1.3 Information islands and copyright protection

Data in cyberspace, especially data from the Internet of things, sensors, and mobile applications, are owned by different persons, companies, and organizations. To protect privacy, ownership, and security, the phenomenon of “information islands” is common in cyberspace big data. Data and knowledge can be hard to gather in a single data center for unified storage, search, analysis, and mining [8]. Experts from various research institutions and organizations around the world have made fruitful achievements in the fields of data analysis, processing, and mining. The technical problems involved are hard for a single company or organization to solve, making it important to bring skills and resources together to solve problems while protecting intellectual property and the ownership of data.

To solve these difficulties, this paper proposes a computing architecture named “fogcloud computing” for big data in ubiquitous cyberspace. It implements collaborative computing of multiple knowledge actors in the fog layer, middle layer, and cloud layer based on collaborative computing language and models. In this paper, a “knowledge actor” is defined as an autonomous, intelligent software agent. It contains a knowledge base, a learning unit, an inference engine, a task executor, and a cooperater. Knowledge actors can not only learn knowledge according to their objectives, but can also work together to complete a task according to the task’s needs. The “fogcloud computing” architecture can be adapted to the Internet of things, intelligent home furnishing, and question-answering robots. “Big search” in cyberspace [9] are typical applications of “fogcloud computing.” The iOS system’s Siri [10] and HiVoice in Huawei phone systems [11] can compute in coordination with local APPs, private data, and Internet data to give intelligent solutions for users. As such, they reflect the basic ideas of “fogcloud computing.”

2 Basic concepts of fogcloud computing

Fogcloud computing can achieve intended goals through the collaborative computing of knowledge actors in the edge layer, middle layer, and cloud center. Fogcloud computing consists of three parts listed in Equation (1): knowledge actors, the relations between knowledge actors, and the operations between knowledge actors.

$$\text{FogcloudComputing} = \{\text{KnowledgeActor}, R, \text{Operation}\} \quad (1)$$

Here, *KnowledgeActor* is the set of all knowledge actors and *KnowledgeActor_i* ($i = 1, \dots, n$) denotes a specified knowledge actor. *R* is the set of relations between knowledge actors, and *R_{ik}* is the relation between *KnowledgeActor_j* and *KnowledgeActor_k*, where $j, k = 1, \dots, n$. *Operation* is the set of operations between knowledge actors; *Operation_m* is the operation associated with *Task_m*.

A knowledge actor is an intelligent agent, varying from simple agents for extracting knowledge from texts to a very complex agent for weather forecasting and other such purposes. Knowledge actors have the following characteristics: self-learning, self-evolution, describability, managability, flexible online composition, and distributed deployment. Knowledge actors can finish a task through coordinated computing among themselves. A knowledge actor *KnowledgeActor_i* = {*ID*, *O*, *A*} consists of Identification, Objective, and Architecture. The objective *O* can be represented as Equation (2):

$$O = \{\text{CollectData}, \text{KnowledgeAcquisition}, \\ \text{Inference}, \text{Computing}, \text{Merging}, \dots\} \quad (2)$$

Objective *O* consists of data collection, knowledge acquisition, knowledge inference, knowledge computing, and knowledge merging.

The architecture of a knowledge actor is shown in Fig. 1. It contains a variety of components, such as a task executor, knowledge base, inference engine, learning unit, task planning and decision maker (TPDK), cooperater, knowledge interface, cooperative interface (CInterface), data interface (DInterface), and task interface (TInterface). The architecture *A* of a knowledge actor can be represented as Equation (3):

$$A = \{\text{TaskExecutor}, \text{KnowledgeBase}, \text{InferEngine}, \\ \text{LearnUnit}, \text{TPDK}, \text{CInterface}, \text{DInterface}, \text{TInterface}\} \quad (3)$$

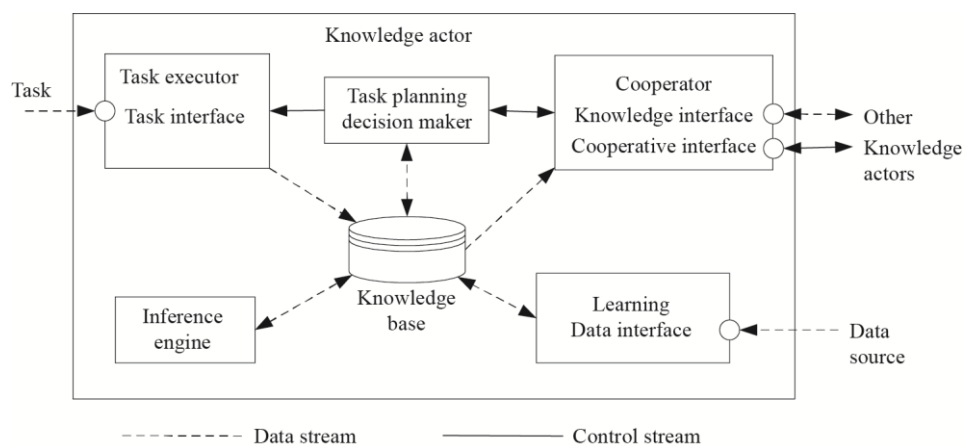


Fig. 1. The architecture of a knowledge actor.

3 The architecture of fogcloud computing

The architecture of fogcloud computing is shown in Fig. 2. It consists of edge computing in the fog layer, fusion computing in the middle layer, and cloud computing in the cloud layer. It is based on the collaborative computing of multiple knowledge actors, and the knowledge actors in the three layers are deployed in different operation platforms. The main difference between fogcloud computing and edge computing or cloud computing is that multiple knowledge actors in the fog, middle, and cloud layers can compute collaboratively, based on the multi-agent architecture.

3.1 Knowledge actors in the cloud layer

Knowledge actors in the cloud layer are in the remote processing centers of fogcloud computing. They gather knowledge from the edge nodes of the Internet of things, Web applications, mobile APPs, and knowledge actors in the other two layers. They are the global center, and are sensitive to contexts and spatiotemporal locations. They usually store global and context data and perform global knowledge inference based on the data.

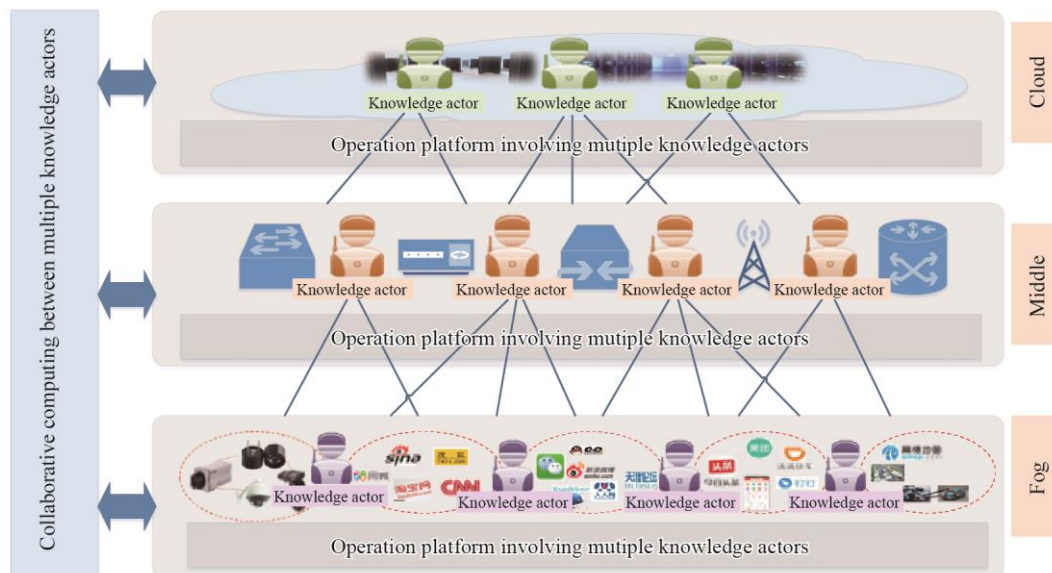


Fig. 2. The architecture of fogcloud computing.

The characteristics of knowledge actors in the cloud layer are as follows: (1) They will have strong computing and storage capacity and very stable networks, (2) they usually carry out global comprehensive analysis and calculation, (3) the knowledge they acquire is widely distributed over large spatiotemporal scales, (4) they perform collaborative computing with knowledge actors in the middle and fog layer, and (5) they usually carry out offline analysis and calculation, owing to the large amount of data they process.

Knowledge actors in the cloud layer are usually based on the following platforms:

(1) *Distributed storage cluster*: The cloud computing platform is oriented to the computing tasks of ubiquitous big data in cyberspace. Knowledge actors in the cloud layer acquire massive data from the Internet, the Internet of things, mobile Internet, and so on. To form knowledge, we need to carry out extracting, processing, merging, and so on. They can carry out specific tasks based on the knowledge. As a platform with a large number of intelligent knowledge actors, it is necessary to store a large number of structured, semi-structured, and unstructured data, including files, web pages, knowledge maps, audios, and videos. The distributed storage cluster is composed of several storage servers and index servers. And it needs specified data storage managements such as the distributed file system and the graph database.

(2) *Artificial intelligence (AI) cluster*: The AI computing cluster is used for high-density computing tasks such as training AI algorithms for pictures, audios, and videos in the knowledge actors in the cloud layer. They are usually in a hybrid CPU + GPU architecture. The cluster's servers are connected with high-speed networks.

(3) *Streaming data processing cluster*: The cluster is used to process massive streaming data from the fog layer or the middle layer for real-time computation. The stream data may come from sensors, mobile terminals, or remote intelligent crawlers. Massive streaming data needs to be cleaned, filtered, and calculated in real-time, and then stored in distributed storage cluster for knowledge discovery, or directly displayed to search users for knowledge discovery.

(4) *High-speed cache cluster*: The cluster usually stores quickly accessed data like the index data and hot data. It mainly uses memory storage to provide support for high real-time computing and the tasks that need to respond quickly.

(5) *Message exchange cluster*: These cluster provides the underlying communication infrastructure for collaborative computing among a large number of knowledge actors. The servers in these clusters are usually connected with high-speed networks.

3.2 Knowledge actors in the middle layer

The knowledge actors in the middle layer lie between the fog layer and the cloud layer. They are usually deployed on servers that are close to the edge nodes. Sometimes they are deployed on the aggregation nodes such as gateways, routers, and communication stations. The knowledge actors in the middle layer usually run on the servers with stronger computing and storage capacity than the servers in the fog layer. Compared with the nodes in the cloud layer, it is closer to the nodes and data source in the fog layer. It is the center where the comprehensive calculations are carried out with the local knowledge actors in the fog layer. The knowledge actors in the middle layer play a key role in the fogcloud computing: on the one hand, the knowledge actors in the middle layer can access the massive data on the Internet and perform collaborative computing with the knowledge actors in the cloud layer; on the other hand, the knowledge actors in the middle layer can access a large number of data and knowledge in multiple local edge nodes (these data and knowledge are information which is usually difficult to quickly return to the cloud center. For example, the user's privacy information usually stores locally, and real-time video monitoring data is too large to transmit to the cloud center in real time). The knowledge actors in the middle layer can also carry out collaborative computing considering the geographical locations of the nodes in the fog layer. They are the combination and key of the global and local computing of data in the cloud and the fog layer. For example, the knowledge actors of smart homes in the middle layer store a large number of home sensor data, video monitoring data, and mobile application data. In order to provide real-time comprehensive feedback and protect home privacy, the local knowledge actors in the middle layer of the fogcloud computing will perform fusion calculation and real-time feedback on the data in the local and remote cloud.

The characteristics of the knowledge actors in the middle layer are as follows: their computing and storage capacity must be strong, and their networks must be stable. They are close to the data sources and can access the original data and perform spatiotemporal analysis and fusion computing. Furthermore, to realize the collaborative reasoning and calculation of multiple knowledge actors (such as video monitoring data and privacy data) and knowledge actors in the cloud, the connected knowledge actors in the fog layer are heterogeneous. This can pose a severe challenge to fusion analysis and calculation. They can carry out near real-time computation, analysis, calculation, and feedback.

3.3 Knowledge actors in the fog layer

The numerous fog edge nodes are the direct sources of data in ubiquitous cyberspace. They include devices in the Internet of things, sensors in sensor networks, APPs in the mobile Internet, smart home devices, and intelligent transportation tools (such as cars and drones installed with intelligent devices). The computing ability and storage capability in the fog layers are relatively small. These knowledge actors are usually for collecting data and carrying out real-time computing.

The characteristics of the knowledge actors in the fog layer are as follows: They have weak computing and storage capacity, often with unstable networks. They are the source of the data in the fogcloud computing and they collect data from the edge nodes. The data they collect contains time and location information. They can move from one place to another. Finally, they usually carry out real-time computation.

3.4 Collaborative computing between multiple knowledge actors

The knowledge actors in the fog layer, middle layer, and cloud layer can carry out collaborative computing and task scheduling based on the multi-agent computing language. The multi-agent computing language defines domains, knowledge actors, actions, problems, and goals. The task scheduling of the multiple knowledge actors is conducted by the task scheduler for multi-objective optimization. After the creation and assembly of knowledge actors, the specific collaborative strategies of multi knowledge actors are generated based on the collaborative task planner. For the collaborative computing tasks of multi knowledge actors, the optimizers with multi-objectives solve the optimized task planning in the space of the task. According to the generated planning, the knowledge actors perform online assembly. The collaborative computing is not only carried out with knowledge actors in the same layer, but also between all three layers of the fogcloud computing architecture. This collaboration between multiple knowledge actors is dynamic. During task execution, knowledge actors can dynamically join or exit a task.

4 Conclusion

This paper presents a software architecture of big data in cyberspace named “Fogcloud Computing” for the collaborative computing of multiple knowledge actors in the fog layer, middle layer, and cloud layer. The architecture is based on multi-agent collaborative computing languages and models. This paper describes the challenges of big data in cyberspace and the basic concepts, and architecture of the fogcloud computing. The key technologies in fogcloud computing include multi-level data representation, multi-agent collaborative reasoning, spatial-temporal fusion analysis, trusted search, and privacy preservation.

The software architecture of fogcloud computing is suitable for typical applications of big data computing in ubiquitous cyberspace, and provides solutions for collaborative computing in the fog, middle, and cloud layer of ubiquitous cyberspace. The fogcloud computing architecture is adaptive to the applications in the Internet of things, smart homes, question-answering robots, and other such sophisticated applications. “Big search” in cyberspace is a typical application of the fogcloud computing. Based on the understanding of users’ intention and the knowledge obtained from the ubiquitous big data, “big search” in cyberspace gives an intelligent solution to meet users’ needs.

References

- [1] Schroeder B, Gibson G. A large-scale study of failures in high performance computing systems [J]. *IEEE Transactions on Dependable and Secure Computing*, 2009, 7(4): 337–350.
- [2] Luo J Z, Jin J H, Song A B, et al. Cloud computing: Architecture and key technologies [J]. *Journal on Communications*, 2011 (7): 3–21. Chinese.
- [3] Bonomi F, Milito R, Zhu J, et al. Fog computing and its role in the internet of things [C]. Helsinki: Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, 2012.
- [4] Shi W, Cao J, Zhang Q, et al. Edge computing: Vision and challenges [J]. *IEEE Internet of Things Journal*, 2016, 3(5): 637–646.
- [5] Liu D Y, Yang K, Chen J Z. Agents: Present status and trends [J]. *Journal of Software*, 2000, 11(3): 315–321. Chinese.
- [6] Olfati-Saber R, Fax J A, Murray R M. Consensus and cooperation in networked multi-agent systems [J]. *Proceedings of the IEEE*, 2007, 95(1): 215–233.
- [7] Mayer-Schonberger V, Cukier K. Big data: A revolution that will transform how we live, work, and think [M]. Boston: Houghton Mifflin Harcourt, 2013.
- [8] Sicari S, Rizzardi A, Grieco L A, et al. Security, privacy and trust in Internet of Things: The road ahead [J]. *Computer Networks*, 2015, 76(15): 146–164.
- [9] Fang B X, Jia Y, Li X, et al. Big search in cyberspace [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2017, 29(9): 1793–1805.
- [10] López G, Quesada L, Guerrero L A. Alexa vs. Siri vs. Cortana vs. Google Assistant: A comparison of speech-based natural user interfaces [C]. Los Angeles: International Conference on Applied Human Factors and Ergonomics, 2017.
- [11] Baker M. Real, unreal, and hacked [J]. *IEEE Pervasive Computing*, 2018, 17(1): 104–112.